

# CMedium\_30\_3\_5\_distilbert

August 10, 2022

## 0.1 Install packages

```
[ ]: !git clone https://github.com/megagonlabs/ditto
      %cd ditto
      !pip install -r requirements.txt
      !git clone https://github.com/rit-git/Snipnext_public
      %cd Snipnext_public
      !pip install -r requirements.txt
      %cd ..

      import nltk
      nltk.download('stopwords')
```

```
Cloning into 'ditto'...
remote: Enumerating objects: 291, done.
remote: Total 291 (delta 0), reused 0 (delta 0), pack-reused 291
Receiving objects: 100% (291/291), 26.87 MiB | 23.26 MiB/s, done.
Resolving deltas: 100% (142/142), done.
/content/ditto
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting gensim==3.8.1
  Downloading gensim-3.8.1-cp37-cp37m-manylinux1_x86_64.whl (24.2 MB)
    |                               | 24.2 MB 1.5 MB/s
Collecting numpy==1.19.2
  Downloading numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
    |                               | 14.5 MB 41.7 MB/s
Collecting regex==2019.12.20
  Downloading regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
    |                               | 689 kB 69.6 MB/s
Collecting scipy==1.3.2
  Downloading scipy-1.3.2-cp37-cp37m-manylinux1_x86_64.whl (25.2 MB)
    |                               | 25.2 MB 1.3 MB/s
Collecting sentencepiece==0.1.85
  Downloading sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
    |                               | 1.0 MB 53.4 MB/s
Requirement already satisfied: sklearn==0.0 in
/usr/local/lib/python3.7/dist-packages (from -r requirements.txt (line 6)) (0.0)
```

```

Collecting spacy==3.1
  Downloading
spacy-3.1.0-cp37-cp37m-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (6.4 MB)
|                                     | 6.4 MB 13.8 MB/s
ERROR: Could not find a version that satisfies the requirement
torch==1.9.0+cu111 (from versions: 1.0.0, 1.0.1, 1.0.1.post2, 1.1.0, 1.2.0,
1.3.0, 1.3.1, 1.4.0, 1.5.0, 1.5.1, 1.6.0, 1.7.0, 1.7.1, 1.8.0, 1.8.1, 1.9.0,
1.9.1, 1.10.0, 1.10.1, 1.10.2, 1.11.0, 1.12.0, 1.12.1)
ERROR: No matching distribution found for torch==1.9.0+cu111
Cloning into 'Snippext_public'...
remote: Enumerating objects: 413, done.
remote: Counting objects: 100% (22/22), done.
remote: Compressing objects: 100% (8/8), done.
remote: Total 413 (delta 15), reused 14 (delta 14), pack-reused 391
Receiving objects: 100% (413/413), 21.06 MiB | 22.25 MiB/s, done.
Resolving deltas: 100% (196/196), done.
/content/ditto/Snippext_public
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting gensim==3.8.1
  Using cached gensim-3.8.1-cp37-cp37m-manylinux1_x86_64.whl (24.2 MB)
Collecting numpy==1.19.2
  Using cached numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
Collecting regex==2019.12.20
  Using cached regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
Collecting spacy==2.2.3
  Downloading spacy-2.2.3-cp37-cp37m-manylinux1_x86_64.whl (10.4 MB)
|                                     | 10.4 MB 4.6 MB/s
Collecting sentencepiece==0.1.85
  Using cached sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
Requirement already satisfied: sklearn==0.0 in /usr/local/lib/python3.7/dist-
packages (from -r requirements.txt (line 6)) (0.0)
Collecting tensorboardX==2.0
  Downloading tensorboardX-2.0-py2.py3-none-any.whl (195 kB)
|                                     | 195 kB 63.3 MB/s
Collecting torch==1.4.0
  Downloading torch-1.4.0-cp37-cp37m-manylinux1_x86_64.whl (753.4 MB)
|                                     | 753.4 MB 7.2 kB/s
Collecting tqdm==4.41.0
  Downloading tqdm-4.41.0-py2.py3-none-any.whl (56 kB)
|                                     | 56 kB 5.1 MB/s
Collecting transformers==3.1.0
  Downloading transformers-3.1.0-py3-none-any.whl (884 kB)
|                                     | 884 kB 67.3 MB/s
Collecting jsonlines==1.2.0
  Downloading jsonlines-1.2.0-py2.py3-none-any.whl (7.6 kB)

```

```

Collecting nltk==3.4.5
  Downloading nltk-3.4.5.zip (1.5 MB)
    | 1.5 MB 63.3 MB/s
Requirement already satisfied: smart-open>=1.8.1 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (5.2.1)
Requirement already satisfied: scipy>=0.18.1 in /usr/local/lib/python3.7/dist-
packages (from gensim==3.8.1->-r requirements.txt (line 1)) (1.7.3)
Requirement already satisfied: six>=1.5.0 in /usr/local/lib/python3.7/dist-
packages (from gensim==3.8.1->-r requirements.txt (line 1)) (1.15.0)
Collecting thinc<7.4.0,>=7.3.0
  Downloading thinc-7.3.1-cp37-cp37m-manylinux1_x86_64.whl (2.2 MB)
    | 2.2 MB 46.3 MB/s
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (3.0.6)
Collecting catalogue<1.1.0,>=0.0.7
  Downloading catalogue-1.0.0-py2.py3-none-any.whl (7.7 kB)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.23.0)
Requirement already satisfied: wasabi<1.1.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (0.10.1)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (1.0.7)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.0.6)
Requirement already satisfied: setuptools in /usr/local/lib/python3.7/dist-
packages (from spacy==2.2.3->-r requirements.txt (line 4)) (57.4.0)
Collecting plac<1.2.0,>=0.9.6
  Downloading plac-1.1.3-py2.py3-none-any.whl (20 kB)
Collecting srsly<1.1.0,>=0.1.0
  Downloading srsly-1.0.5-cp37-cp37m-manylinux2014_x86_64.whl (184 kB)
    | 184 kB 57.9 MB/s
Collecting blis<0.5.0,>=0.4.0
  Downloading blis-0.4.1-cp37-cp37m-manylinux1_x86_64.whl (3.7 MB)
    | 3.7 MB 44.8 MB/s
Requirement already satisfied: scikit-learn in
/usr/local/lib/python3.7/dist-packages (from sklearn==0.0->-r requirements.txt
(line 6)) (1.0.2)
Requirement already satisfied: protobuf>=3.8.0 in /usr/local/lib/python3.7/dist-
packages (from tensorboardX==2.0->-r requirements.txt (line 8)) (3.17.3)
Collecting sacremoses
  Downloading sacremoses-0.0.53.tar.gz (880 kB)
    | 880 kB 62.5 MB/s

```

```

Collecting tokenizers==0.8.1.rc2
  Downloading tokenizers-0.8.1rc2-cp37-cp37m-manylinux1_x86_64.whl (3.0 MB)
    |                                     | 3.0 MB 55.4 MB/s
Requirement already satisfied: filelock in /usr/local/lib/python3.7/dist-packages (from transformers==3.1.0->-r requirements.txt (line 11)) (3.7.1)
Requirement already satisfied: packaging in /usr/local/lib/python3.7/dist-packages (from transformers==3.1.0->-r requirements.txt (line 11)) (21.3)
Requirement already satisfied: importlib-metadata>=0.20 in /usr/local/lib/python3.7/dist-packages (from catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line 4)) (4.12.0)
Requirement already satisfied: typing-extensions>=3.6.4 in /usr/local/lib/python3.7/dist-packages (from importlib-metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line 4)) (4.1.1)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-packages (from importlib-metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line 4)) (3.8.1)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-packages (from requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (2.10)
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in /usr/local/lib/python3.7/dist-packages (from requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (1.24.3)
Requirement already satisfied: chardet<4,>=3.0.2 in /usr/local/lib/python3.7/dist-packages (from requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (3.0.4)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.7/dist-packages (from requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (2022.6.15)
Requirement already satisfied: pyparsing!=3.0.5,>=2.0.2 in /usr/local/lib/python3.7/dist-packages (from packaging->transformers==3.1.0->-r requirements.txt (line 11)) (3.0.9)
Requirement already satisfied: click in /usr/local/lib/python3.7/dist-packages (from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (7.1.2)
Requirement already satisfied: joblib in /usr/local/lib/python3.7/dist-packages (from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (1.1.0)
Requirement already satisfied: threadpoolctl>=2.0.0 in /usr/local/lib/python3.7/dist-packages (from scikit-learn->sklearn==0.0->-r requirements.txt (line 6)) (3.1.0)
Building wheels for collected packages: nltk, sacremoses
  Building wheel for nltk (setup.py) ... done
  Created wheel for nltk: filename=nltk-3.4.5-py3-none-any.whl size=1449922 sha256=a4d77a16032a91c611033d4ff4635bdc39f5aadd09d3c7d276fdb210f606169
  Stored in directory: /root/.cache/pip/wheels/48/8b/7f/473521e0c731c6566d631b281f323842bbda9bd819eb9a3ead
  Building wheel for sacremoses (setup.py) ... done
  Created wheel for sacremoses: filename=sacremoses-0.0.53-py3-none-any.whl

```

```

size=895260
sha256=1d5e8dced5cc2e449929342dc46796dbe102851d1630bf2846c658363cc4f005
  Stored in directory: /root/.cache/pip/wheels/87/39/dd/a83eeef36d0bf98e7a4d1933
a4ad2d660295a40613079bafc9
Successfully built nltk sacremoses
Installing collected packages: numpy, tqdm, srsly, regex, plac, blis,
tokenizers, thinc, sentencepiece, sacremoses, catalogue, transformers, torch,
tensorboardX, spacy, nltk, jsonlines, gensim
  Attempting uninstall: numpy
    Found existing installation: numpy 1.21.6
    Uninstalling numpy-1.21.6:
      Successfully uninstalled numpy-1.21.6
  Attempting uninstall: tqdm
    Found existing installation: tqdm 4.64.0
    Uninstalling tqdm-4.64.0:
      Successfully uninstalled tqdm-4.64.0
  Attempting uninstall: srsly
    Found existing installation: srsly 2.4.4
    Uninstalling srsly-2.4.4:
      Successfully uninstalled srsly-2.4.4
  Attempting uninstall: regex
    Found existing installation: regex 2022.6.2
    Uninstalling regex-2022.6.2:
      Successfully uninstalled regex-2022.6.2
  Attempting uninstall: blis
    Found existing installation: blis 0.7.8
    Uninstalling blis-0.7.8:
      Successfully uninstalled blis-0.7.8
  Attempting uninstall: thinc
    Found existing installation: thinc 8.1.0
    Uninstalling thinc-8.1.0:
      Successfully uninstalled thinc-8.1.0
  Attempting uninstall: catalogue
    Found existing installation: catalogue 2.0.8
    Uninstalling catalogue-2.0.8:
      Successfully uninstalled catalogue-2.0.8
  Attempting uninstall: torch
    Found existing installation: torch 1.12.0+cu113
    Uninstalling torch-1.12.0+cu113:
      Successfully uninstalled torch-1.12.0+cu113
  Attempting uninstall: spacy
    Found existing installation: spacy 3.4.1
    Uninstalling spacy-3.4.1:
      Successfully uninstalled spacy-3.4.1
  Attempting uninstall: nltk
    Found existing installation: nltk 3.7
    Uninstalling nltk-3.7:
      Successfully uninstalled nltk-3.7

```

```
Attempting uninstall: gensim
Found existing installation: gensim 3.6.0
Uninstalling gensim-3.6.0:
Successfully uninstalled gensim-3.6.0
```

ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behaviour is the source of the following dependency conflicts.

xarray-einstats 0.2.2 requires numpy>=1.21, but you have numpy 1.19.2 which is incompatible.

torchvision 0.13.0+cu113 requires torch==1.12.0, but you have torch 1.4.0 which is incompatible.

torchtext 0.13.0 requires torch==1.12.0, but you have torch 1.4.0 which is incompatible.

torchaudio 0.12.0+cu113 requires torch==1.12.0, but you have torch 1.4.0 which is incompatible.

tensorflow 2.8.2+zzzcolab20220719082949 requires numpy>=1.20, but you have numpy 1.19.2 which is incompatible.

panel 0.12.1 requires tqdm>=4.48.0, but you have tqdm 4.41.0 which is incompatible.

fastai 2.7.7 requires torch<1.13,>=1.7, but you have torch 1.4.0 which is incompatible.

en-core-web-sm 3.4.0 requires spacy<3.5.0,>=3.4.0, but you have spacy 2.2.3 which is incompatible.

cmdstanpy 1.0.4 requires numpy>=1.21, but you have numpy 1.19.2 which is incompatible.

```
Successfully installed blis-0.4.1 catalogue-1.0.0 gensim-3.8.1 jsonlines-1.2.0
nltk-3.4.5 numpy-1.19.2 plac-1.1.3 regex-2019.12.20 sacremoses-0.0.53
sentencepiece-0.1.85 spacy-2.2.3 srsly-1.0.5 tensorboardX-2.0 thinc-7.3.1
tokenizers-0.8.1rc2 torch-1.4.0 tqdm-4.41.0 transformers-3.1.0
```

```
/content/ditto
```

```
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data] Unzipping corpora/stopwords.zip.
```

```
[ ]: True
```

## 0.2 Install fp16 optimization

```
[ ]: !git clone https://github.com/NVIDIA/apex
      %cd apex
      !pip install -v --no-cache-dir ./
      %cd ..
```

```
Cloning into 'apex'...
remote: Enumerating objects: 10085, done.
remote: Counting objects: 100% (201/201), done.
remote: Compressing objects: 100% (139/139), done.
remote: Total 10085 (delta 98), reused 130 (delta 58), pack-reused 9884
Receiving objects: 100% (10085/10085), 14.95 MiB | 23.73 MiB/s, done.
Resolving deltas: 100% (6903/6903), done.
/content/ditto/apex
Using pip 21.1.3 from /usr/local/lib/python3.7/dist-packages/pip (python 3.7)
Value for scheme.platlib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.purelib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/UNKNOWN
sysconfig: /usr/include/python3.7m/UNKNOWN
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None
Non-user install because site-packages writable
Created temporary directory: /tmp/pip-ephem-wheel-cache-pmttgsbe
Created temporary directory: /tmp/pip-req-tracker-h_el88v4
Initialized build tracking at /tmp/pip-req-tracker-h_el88v4
Created build tracker: /tmp/pip-req-tracker-h_el88v4
Entered build tracker: /tmp/pip-req-tracker-h_el88v4
```

```

Created temporary directory: /tmp/pip-install-1zy0ujn0
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Processing /content/ditto/apex
  Created temporary directory: /tmp/pip-req-build-0ooc51qq
  DEPRECATION: A future pip version will change local packages to be built
in-place without first copying to a temporary directory. We recommend you use
--use-feature=in-tree-build to test your packages with this new behavior before
it becomes the default.

  pip 21.3 will remove support for this functionality. You can find discussion
regarding this at https://github.com/pypa/pip/issues/7555.
  Added file:///content/ditto/apex to build tracker '/tmp/pip-req-tracker-
h_el88v4'
  Running setup.py (path:/tmp/pip-req-build-0ooc51qq/setup.py) egg_info for
package from file:///content/ditto/apex
  Created temporary directory: /tmp/pip-pip-egg-info-3mim56aq
  Running command python setup.py egg_info

  torch.__version__ = 1.4.0

  running egg_info
  creating /tmp/pip-pip-egg-info-3mim56aq/apex.egg-info
  writing /tmp/pip-pip-egg-info-3mim56aq/apex.egg-info/PKG-INFO
  writing dependency_links to /tmp/pip-pip-egg-info-3mim56aq/apex.egg-
info/dependency_links.txt
  writing top-level names to /tmp/pip-pip-egg-info-3mim56aq/apex.egg-
info/top_level.txt
  writing manifest file '/tmp/pip-pip-egg-info-3mim56aq/apex.egg-
info/SOURCES.txt'
  adding license file 'LICENSE'
  writing manifest file '/tmp/pip-pip-egg-info-3mim56aq/apex.egg-
info/SOURCES.txt'
  Source in /tmp/pip-req-build-0ooc51qq has version 0.1, which satisfies
requirement apex==0.1 from file:///content/ditto/apex
  Removed apex==0.1 from file:///content/ditto/apex from build tracker
'/tmp/pip-req-tracker-h_el88v4'
Created temporary directory: /tmp/pip-unpack-9vgnzc7z
Building wheels for collected packages: apex
  Created temporary directory: /tmp/pip-wheel-eyfr_ha1
  Building wheel for apex (setup.py) ... Destination directory:
/tmp/pip-wheel-eyfr_ha1
  Running command /usr/bin/python3 -u -c 'import io, os, sys, setuptools,
tokenize; sys.argv[0] = ''"/tmp/pip-req-build-0ooc51qq/setup.py''';

```



```
__file__='''/tmp/pip-req-build-0ooc51qq/setup.py''';f = getattr(tokenize,
''''open'''' , open)(__file__) if os.path.exists(__file__) else
io.StringIO('''from setuptools import setup; setup()''');code =
f.read().replace('''\r\n''', '''\n''');f.close();exec(compile(code,
__file__, '''exec'''))' bdist_wheel -d /tmp/pip-wheel-eyfr_ha1
```

```
torch.__version__ = 1.4.0
```

```
running bdist_wheel
running build
running build_py
creating build
creating build/lib
creating build/lib/apex
copying apex/__init__.py -> build/lib/apex
copying apex/_autocast_utils.py -> build/lib/apex
creating build/lib/apex/normalization
copying apex/normalization/__init__.py -> build/lib/apex/normalization
copying apex/normalization/fused_layer_norm.py -> build/lib/apex/normalization
creating build/lib/apex/RNN
copying apex/RNN/__init__.py -> build/lib/apex/RNN
copying apex/RNN/cells.py -> build/lib/apex/RNN
copying apex/RNN/RNNBackend.py -> build/lib/apex/RNN
copying apex/RNN/models.py -> build/lib/apex/RNN
creating build/lib/apex/contrib
copying apex/contrib/__init__.py -> build/lib/apex/contrib
creating build/lib/apex/multi_tensor_apply
copying apex/multi_tensor_apply/__init__.py ->
build/lib/apex/multi_tensor_apply
copying apex/multi_tensor_apply/multi_tensor_apply.py ->
build/lib/apex/multi_tensor_apply
creating build/lib/apex/transformer
copying apex/transformer/__init__.py -> build/lib/apex/transformer
copying apex/transformer/enums.py -> build/lib/apex/transformer
copying apex/transformer/microbatches.py -> build/lib/apex/transformer
copying apex/transformer/log_util.py -> build/lib/apex/transformer
copying apex/transformer/utils.py -> build/lib/apex/transformer
copying apex/transformer/parallel_state.py -> build/lib/apex/transformer
creating build/lib/apex/amp
copying apex/amp/__init__.py -> build/lib/apex/amp
copying apex/amp/opt.py -> build/lib/apex/amp
copying apex/amp/rnn_compat.py -> build/lib/apex/amp
copying apex/amp/_process_optimizer.py -> build/lib/apex/amp
copying apex/amp/compat.py -> build/lib/apex/amp
copying apex/amp/__version__.py -> build/lib/apex/amp
copying apex/amp/utils.py -> build/lib/apex/amp
```

```

copying apex/amp/scaler.py -> build/lib/apex/amp
copying apex/amp/wrap.py -> build/lib/apex/amp
copying apex/amp/frontend.py -> build/lib/apex/amp
copying apex/amp/_initialize.py -> build/lib/apex/amp
copying apex/amp/handle.py -> build/lib/apex/amp
copying apex/amp/_amp_state.py -> build/lib/apex/amp
copying apex/amp/amp.py -> build/lib/apex/amp
creating build/lib/apex/optimizers
copying apex/optimizers/__init__.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_sgd.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_mixed_precision_lamb.py ->
build/lib/apex/optimizers
copying apex/optimizers/fused_novograd.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_adagrad.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_adam.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_lamb.py -> build/lib/apex/optimizers
creating build/lib/apex/fp16_utils
copying apex/fp16_utils/__init__.py -> build/lib/apex/fp16_utils
copying apex/fp16_utils/loss_scaler.py -> build/lib/apex/fp16_utils
copying apex/fp16_utils/fp16util.py -> build/lib/apex/fp16_utils
copying apex/fp16_utils/fp16_optimizer.py -> build/lib/apex/fp16_utils
creating build/lib/apex/fused_dense
copying apex/fused_dense/__init__.py -> build/lib/apex/fused_dense
copying apex/fused_dense/fused_dense.py -> build/lib/apex/fused_dense
creating build/lib/apex/parallel
copying apex/parallel/optimized_sync_batchnorm.py -> build/lib/apex/parallel
copying apex/parallel/__init__.py -> build/lib/apex/parallel
copying apex/parallel/optimized_sync_batchnorm_kernel.py ->
build/lib/apex/parallel
copying apex/parallel/sync_batchnorm_kernel.py -> build/lib/apex/parallel
copying apex/parallel/multiproc.py -> build/lib/apex/parallel
copying apex/parallel/distributed.py -> build/lib/apex/parallel
copying apex/parallel/LARC.py -> build/lib/apex/parallel
copying apex/parallel/sync_batchnorm.py -> build/lib/apex/parallel
creating build/lib/apex/mlp
copying apex/mlp/__init__.py -> build/lib/apex/mlp
copying apex/mlp/mlp.py -> build/lib/apex/mlp
creating build/lib/apex/contrib/focal_loss
copying apex/contrib/focal_loss/__init__.py ->
build/lib/apex/contrib/focal_loss
copying apex/contrib/focal_loss/focal_loss.py ->
build/lib/apex/contrib/focal_loss
creating build/lib/apex/contrib/peer_memory
copying apex/contrib/peer_memory/__init__.py ->
build/lib/apex/contrib/peer_memory
copying apex/contrib/peer_memory/peer_halo_exchanger_1d.py ->
build/lib/apex/contrib/peer_memory
copying apex/contrib/peer_memory/peer_memory.py ->

```

```

build/lib/apex/contrib/peer_memory
  copying apex/contrib/peer_memory/peer_halo_exchange_module_tests.py ->
build/lib/apex/contrib/peer_memory
  creating build/lib/apex/contrib/groupbn
  copying apex/contrib/groupbn/__init__.py -> build/lib/apex/contrib/groupbn
  copying apex/contrib/groupbn/batch_norm.py -> build/lib/apex/contrib/groupbn
  creating build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/__init__.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/distributed_fused_adam.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/fused_sgd.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/fused_adam.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/fused_lamb.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/distributed_fused_lamb.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/fp16_optimizer.py ->
build/lib/apex/contrib/optimizers
  creating build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/__init__.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/self_multihead_attn.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/fast_self_multihead_attn_norm_add_func.py
-> build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/fast_self_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/encdec_multihead_attn.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/fast_encdec_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/encdec_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
  copying
apex/contrib/multihead_attn/fast_encdec_multihead_attn_norm_add_func.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/mask_softmax_dropout_func.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/self_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
  creating build/lib/apex/contrib/bottleneck
  copying apex/contrib/bottleneck/__init__.py ->
build/lib/apex/contrib/bottleneck
  copying apex/contrib/bottleneck/bottleneck_module_test.py ->
build/lib/apex/contrib/bottleneck

```

```

    copying apex/contrib/bottleneck/test.py -> build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/bottleneck.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/halo_exchangers.py ->
build/lib/apex/contrib/bottleneck
    creating build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/__init__.py -> build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/permutation_lib.py ->
build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/asp.py -> build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/sparse_masklib.py ->
build/lib/apex/contrib/sparsity
    creating build/lib/apex/contrib/layer_norm
    copying apex/contrib/layer_norm/__init__.py ->
build/lib/apex/contrib/layer_norm
    copying apex/contrib/layer_norm/layer_norm.py ->
build/lib/apex/contrib/layer_norm
    creating build/lib/apex/contrib/transducer
    copying apex/contrib/transducer/__init__.py ->
build/lib/apex/contrib/transducer
    copying apex/contrib/transducer/transducer.py ->
build/lib/apex/contrib/transducer
    creating build/lib/apex/contrib/index_mul_2d
    copying apex/contrib/index_mul_2d/__init__.py ->
build/lib/apex/contrib/index_mul_2d
    copying apex/contrib/index_mul_2d/index_mul_2d.py ->
build/lib/apex/contrib/index_mul_2d
    creating build/lib/apex/contrib/conv_bias_relu
    copying apex/contrib/conv_bias_relu/__init__.py ->
build/lib/apex/contrib/conv_bias_relu
    copying apex/contrib/conv_bias_relu/conv_bias_relu.py ->
build/lib/apex/contrib/conv_bias_relu
    creating build/lib/apex/contrib/xentropy
    copying apex/contrib/xentropy/__init__.py -> build/lib/apex/contrib/xentropy
    copying apex/contrib/xentropy/softmax_xentropy.py ->
build/lib/apex/contrib/xentropy
    creating build/lib/apex/contrib/clip_grad
    copying apex/contrib/clip_grad/__init__.py -> build/lib/apex/contrib/clip_grad
    copying apex/contrib/clip_grad/clip_grad.py ->
build/lib/apex/contrib/clip_grad
    creating build/lib/apex/contrib/fmha
    copying apex/contrib/fmha/__init__.py -> build/lib/apex/contrib/fmha
    copying apex/contrib/fmha/fmha.py -> build/lib/apex/contrib/fmha
    creating build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying apex/contrib/sparsity/permutation_search_kernels/__init__.py ->
build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying apex/contrib/sparsity/permutation_search_kernels/call_permutation_sear
ch_kernels.py -> build/lib/apex/contrib/sparsity/permutation_search_kernels

```

```

    copying
apex/contrib/sparsity/permutation_search_kernels/permutation_utilities.py ->
build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying apex/contrib/sparsity/permutation_search_kernels/exhaustive_search.py
-> build/lib/apex/contrib/sparsity/permutation_search_kernels
    creating build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/__init__.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/mappings.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/layers.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/memory.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/utils.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/cross_entropy.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/random.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/data.py ->
build/lib/apex/transformer/tensor_parallel
    creating build/lib/apex/transformer/testing
    copying apex/transformer/testing/__init__.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/distributed_test_base.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/standalone_bert.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/arguments.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/standalone_transformer_lm.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/commons.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/global_vars.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/standalone_gpt.py ->
build/lib/apex/transformer/testing
    creating build/lib/apex/transformer/functional
    copying apex/transformer/functional/__init__.py ->
build/lib/apex/transformer/functional
    copying apex/transformer/functional/fused_softmax.py ->
build/lib/apex/transformer/functional
    creating build/lib/apex/transformer/amp
    copying apex/transformer/amp/__init__.py -> build/lib/apex/transformer/amp
    copying apex/transformer/amp/grad_scaler.py -> build/lib/apex/transformer/amp
    creating build/lib/apex/transformer/_data

```

```

    copying apex/transformer/_data/__init__.py -> build/lib/apex/transformer/_data
    copying apex/transformer/_data/_batch_sampler.py ->
build/lib/apex/transformer/_data
    creating build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/__init__.py ->
build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/_timers.py ->
build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/utils.py ->
build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/p2p_communication.py ->
build/lib/apex/transformer/pipeline_parallel
    creating build/lib/apex/transformer/layers
    copying apex/transformer/layers/__init__.py ->
build/lib/apex/transformer/layers
    copying apex/transformer/layers/layer_norm.py ->
build/lib/apex/transformer/layers
    creating build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipeline_parallel_with_interleaving.py -> build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/__init__.py ->
build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_no_pipeline_parallel.py -> build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/common.py ->
build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipeline_parallel_without_interleaving.py -> build/lib/apex/transformer/pipeline_parallel/schedules
    creating build/lib/apex/amp/lists
    copying apex/amp/lists/__init__.py -> build/lib/apex/amp/lists
    copying apex/amp/lists/tensor_overrides.py -> build/lib/apex/amp/lists
    copying apex/amp/lists/functional_overrides.py -> build/lib/apex/amp/lists
    copying apex/amp/lists/torch_overrides.py -> build/lib/apex/amp/lists
    installing to build/bdist.linux-x86_64/wheel
    running install
    running install_lib
    creating build/bdist.linux-x86_64
    creating build/bdist.linux-x86_64/wheel
    creating build/bdist.linux-x86_64/wheel/apex
    copying build/lib/apex/__init__.py -> build/bdist.linux-x86_64/wheel/apex
    copying build/lib/apex/_autocast_utils.py ->
build/bdist.linux-x86_64/wheel/apex
    creating build/bdist.linux-x86_64/wheel/apex/normalization
    copying build/lib/apex/normalization/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/normalization
    copying build/lib/apex/normalization/fused_layer_norm.py ->
build/bdist.linux-x86_64/wheel/apex/normalization
    creating build/bdist.linux-x86_64/wheel/apex/RNN

```

```

    copying build/lib/apex/RNN/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/cells.py -> build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/RNNBackend.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/models.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    creating build/bdist.linux-x86_64/wheel/apex/contrib
    creating build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
    copying build/lib/apex/contrib/focal_loss/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
    copying build/lib/apex/contrib/focal_loss/focal_loss.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
    copying build/lib/apex/contrib/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib
    creating build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    copying build/lib/apex/contrib/peer_memory/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    copying build/lib/apex/contrib/peer_memory/peer_halo_exchanger_1d.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    copying build/lib/apex/contrib/peer_memory/peer_memory.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    copying build/lib/apex/contrib/peer_memory/peer_halo_exchange_module_tests.py
-> build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    creating build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
    copying build/lib/apex/contrib/groupbn/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
    copying build/lib/apex/contrib/groupbn/batch_norm.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
    creating build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/distributed_fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/fused_sgd.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/fused_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/distributed_fused_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/fp16_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    creating build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/self_multihead_attn.py ->

```

```

build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying
build/lib/apex/contrib/multihead_attn/fast_self_multihead_attn_norm_add_func.py
-> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/fast_self_multihead_attn_func.py
-> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/encdec_multihead_attn.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying
build/lib/apex/contrib/multihead_attn/fast_encdec_multihead_attn_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/encdec_multihead_attn_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/fast_encdec_multihead_attn_norm_
add_func.py -> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/mask_softmax_dropout_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/self_multihead_attn_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    creating build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    copying build/lib/apex/contrib/bottleneck/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    copying build/lib/apex/contrib/bottleneck/bottleneck_module_test.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    copying build/lib/apex/contrib/bottleneck/test.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    copying build/lib/apex/contrib/bottleneck/bottleneck.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    copying build/lib/apex/contrib/bottleneck/halo_exchangers.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    creating build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/permutation_lib.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/asp.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/sparse_masklib.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    creating
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/permutation_search_kernels/__init__.py
->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/permutation_search_kernels/call_permut
ation_search_kernels.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/permutation_search_kernels/permutation

```



```

_utilities.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
  copying
build/lib/apex/contrib/sparsity/permutation_search_kernels/exhaustive_search.py
->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
  creating build/bdist.linux-x86_64/wheel/apex/contrib/layer_norm
  copying build/lib/apex/contrib/layer_norm/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/layer_norm
  copying build/lib/apex/contrib/layer_norm/layer_norm.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/layer_norm
  creating build/bdist.linux-x86_64/wheel/apex/contrib/transducer
  copying build/lib/apex/contrib/transducer/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/transducer
  copying build/lib/apex/contrib/transducer/transducer.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/transducer
  creating build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
  copying build/lib/apex/contrib/index_mul_2d/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
  copying build/lib/apex/contrib/index_mul_2d/index_mul_2d.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
  creating build/bdist.linux-x86_64/wheel/apex/contrib/conv_bias_relu
  copying build/lib/apex/contrib/conv_bias_relu/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/conv_bias_relu
  copying build/lib/apex/contrib/conv_bias_relu/conv_bias_relu.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/conv_bias_relu
  creating build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
  copying build/lib/apex/contrib/xentropy/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
  copying build/lib/apex/contrib/xentropy/softmax_xentropy.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
  creating build/bdist.linux-x86_64/wheel/apex/contrib/clip_grad
  copying build/lib/apex/contrib/clip_grad/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/clip_grad
  copying build/lib/apex/contrib/clip_grad/clip_grad.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/clip_grad
  creating build/bdist.linux-x86_64/wheel/apex/contrib/fmha
  copying build/lib/apex/contrib/fmha/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/fmha
  copying build/lib/apex/contrib/fmha/fmha.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/fmha
  creating build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
  copying build/lib/apex/multi_tensor_apply/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
  copying build/lib/apex/multi_tensor_apply/multi_tensor_apply.py ->
build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
  creating build/bdist.linux-x86_64/wheel/apex/transformer
  copying build/lib/apex/transformer/__init__.py ->

```

```

build/bdist.linux-x86_64/wheel/apex/transformer
  creating build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/mappings.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/layers.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/memory.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/utils.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/cross_entropy.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/random.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/data.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/enums.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
  copying build/lib/apex/transformer/microbatches.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
  creating build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/testing/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/testing/distributed_test_base.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/testing/standalone_bert.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/testing/arguments.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/testing/standalone_transformer_lm.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/testing/commons.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/testing/global_vars.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/testing/standalone_gpt.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
  creating build/bdist.linux-x86_64/wheel/apex/transformer/functional
  copying build/lib/apex/transformer/functional/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/functional
  copying build/lib/apex/transformer/functional/fused_softmax.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/functional
  copying build/lib/apex/transformer/log_util.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
  creating build/bdist.linux-x86_64/wheel/apex/transformer/amp
  copying build/lib/apex/transformer/amp/__init__.py ->

```

```

build/bdist.linux-x86_64/wheel/apex/transformer/amp
    copying build/lib/apex/transformer/amp/grad_scaler.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/amp
    copying build/lib/apex/transformer/utils.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/transformer/_data
    copying build/lib/apex/transformer/_data/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/_data
    copying build/lib/apex/transformer/_data/_batch_sampler.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/_data
    creating build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    copying build/lib/apex/transformer/pipeline_parallel/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    copying build/lib/apex/transformer/pipeline_parallel/_timers.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    creating
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipeline_parallel_with_interleaving.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/schedules/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying
build/lib/apex/transformer/pipeline_parallel/schedules/fwd_bwd_no_pipeline_parallel.py
-> build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/schedules/common.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipeline_parallel_without_interleaving.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/utils.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    copying build/lib/apex/transformer/pipeline_parallel/p2p_communication.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    creating build/bdist.linux-x86_64/wheel/apex/transformer/layers
    copying build/lib/apex/transformer/layers/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/layers
    copying build/lib/apex/transformer/layers/layer_norm.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/layers
    copying build/lib/apex/transformer/parallel_state.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/opt.py -> build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/rnn_compat.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/_process_optimizer.py ->

```

```

build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/compat.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  creating build/bdist.linux-x86_64/wheel/apex/amp/lists
  copying build/lib/apex/amp/lists/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
  copying build/lib/apex/amp/lists/tensor_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
  copying build/lib/apex/amp/lists/functional_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
  copying build/lib/apex/amp/lists/torch_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
  copying build/lib/apex/amp/__version__.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/utils.py -> build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/scaler.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/wrap.py -> build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/frontend.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/_initialize.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/handle.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/_amp_state.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/amp.py -> build/bdist.linux-x86_64/wheel/apex/amp
  creating build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_sgd.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_mixed_precision_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_novograd.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_adagrad.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  creating build/bdist.linux-x86_64/wheel/apex/fp16_utils
  copying build/lib/apex/fp16_utils/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
  copying build/lib/apex/fp16_utils/loss_scaler.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
  copying build/lib/apex/fp16_utils/fp16util.py ->

```

```

build/bdist.linux-x86_64/wheel/apex/fp16_utils
  copying build/lib/apex/fp16_utils/fp16_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
  creating build/bdist.linux-x86_64/wheel/apex/fused_dense
  copying build/lib/apex/fused_dense/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/fused_dense
  copying build/lib/apex/fused_dense/fused_dense.py ->
build/bdist.linux-x86_64/wheel/apex/fused_dense
  creating build/bdist.linux-x86_64/wheel/apex/parallel
  copying build/lib/apex/parallel/optimized_sync_batchnorm.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
  copying build/lib/apex/parallel/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
  copying build/lib/apex/parallel/optimized_sync_batchnorm_kernel.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
  copying build/lib/apex/parallel/sync_batchnorm_kernel.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
  copying build/lib/apex/parallel/multiproc.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
  copying build/lib/apex/parallel/distributed.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
  copying build/lib/apex/parallel/LARC.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
  copying build/lib/apex/parallel/sync_batchnorm.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
  creating build/bdist.linux-x86_64/wheel/apex/mlp
  copying build/lib/apex/mlp/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/mlp
  copying build/lib/apex/mlp/mlp.py -> build/bdist.linux-x86_64/wheel/apex/mlp
running install_egg_info
running egg_info
creating apex.egg-info
writing apex.egg-info/PKG-INFO
writing dependency_links to apex.egg-info/dependency_links.txt
writing top-level names to apex.egg-info/top_level.txt
writing manifest file 'apex.egg-info/SOURCES.txt'
adding license file 'LICENSE'
writing manifest file 'apex.egg-info/SOURCES.txt'
Copying apex.egg-info to build/bdist.linux-x86_64/wheel/apex-0.1-py3.7.egg-
info
running install_scripts
adding license file "LICENSE" (matched pattern "LICEN[CS]E*")
creating build/bdist.linux-x86_64/wheel/apex-0.1.dist-info/WHEEL
creating '/tmp/pip-wheel-eyfr_ha1/apex-0.1-py3-none-any.whl' and adding
'build/bdist.linux-x86_64/wheel' to it
  adding 'apex/__init__.py'
  adding 'apex/_autocast_utils.py'
  adding 'apex/RNN/RNNBackend.py'

```

```

adding 'apex/RNN/__init__.py'
adding 'apex/RNN/cells.py'
adding 'apex/RNN/models.py'
adding 'apex/amp/__init__.py'
adding 'apex/amp/__version__.py'
adding 'apex/amp/_amp_state.py'
adding 'apex/amp/_initialize.py'
adding 'apex/amp/_process_optimizer.py'
adding 'apex/amp/amp.py'
adding 'apex/amp/compat.py'
adding 'apex/amp/frontend.py'
adding 'apex/amp/handle.py'
adding 'apex/amp/opt.py'
adding 'apex/amp/rnn_compat.py'
adding 'apex/amp/scaler.py'
adding 'apex/amp/utils.py'
adding 'apex/amp/wrap.py'
adding 'apex/amp/lists/__init__.py'
adding 'apex/amp/lists/functional_overrides.py'
adding 'apex/amp/lists/tensor_overrides.py'
adding 'apex/amp/lists/torch_overrides.py'
adding 'apex/contrib/__init__.py'
adding 'apex/contrib/bottleneck/__init__.py'
adding 'apex/contrib/bottleneck/bottleneck.py'
adding 'apex/contrib/bottleneck/bottleneck_module_test.py'
adding 'apex/contrib/bottleneck/halo_exchangers.py'
adding 'apex/contrib/bottleneck/test.py'
adding 'apex/contrib/clip_grad/__init__.py'
adding 'apex/contrib/clip_grad/clip_grad.py'
adding 'apex/contrib/conv_bias_relu/__init__.py'
adding 'apex/contrib/conv_bias_relu/conv_bias_relu.py'
adding 'apex/contrib/fmha/__init__.py'
adding 'apex/contrib/fmha/fmha.py'
adding 'apex/contrib/focal_loss/__init__.py'
adding 'apex/contrib/focal_loss/focal_loss.py'
adding 'apex/contrib/groupbn/__init__.py'
adding 'apex/contrib/groupbn/batch_norm.py'
adding 'apex/contrib/index_mul_2d/__init__.py'
adding 'apex/contrib/index_mul_2d/index_mul_2d.py'
adding 'apex/contrib/layer_norm/__init__.py'
adding 'apex/contrib/layer_norm/layer_norm.py'
adding 'apex/contrib/multihead_attn/__init__.py'
adding 'apex/contrib/multihead_attn/encdec_multihead_attn.py'
adding 'apex/contrib/multihead_attn/encdec_multihead_attn_func.py'
adding 'apex/contrib/multihead_attn/fast_encdec_multihead_attn_func.py'
adding
'apex/contrib/multihead_attn/fast_encdec_multihead_attn_norm_add_func.py'
    adding 'apex/contrib/multihead_attn/fast_self_multihead_attn_func.py'

```

```

adding 'apex/contrib/multihead_attn/fast_self_multihead_attn_norm_add_func.py'
adding 'apex/contrib/multihead_attn/mask_softmax_dropout_func.py'
adding 'apex/contrib/multihead_attn/self_multihead_attn.py'
adding 'apex/contrib/multihead_attn/self_multihead_attn_func.py'
adding 'apex/contrib/optimizers/__init__.py'
adding 'apex/contrib/optimizers/distributed_fused_adam.py'
adding 'apex/contrib/optimizers/distributed_fused_lamb.py'
adding 'apex/contrib/optimizers/fp16_optimizer.py'
adding 'apex/contrib/optimizers/fused_adam.py'
adding 'apex/contrib/optimizers/fused_lamb.py'
adding 'apex/contrib/optimizers/fused_sgd.py'
adding 'apex/contrib/peer_memory/__init__.py'
adding 'apex/contrib/peer_memory/peer_halo_exchange_module_tests.py'
adding 'apex/contrib/peer_memory/peer_halo_exchanger_1d.py'
adding 'apex/contrib/peer_memory/peer_memory.py'
adding 'apex/contrib/sparsity/__init__.py'
adding 'apex/contrib/sparsity/asp.py'
adding 'apex/contrib/sparsity/permutation_lib.py'
adding 'apex/contrib/sparsity/sparse_masklib.py'
adding 'apex/contrib/sparsity/permutation_search_kernels/__init__.py'
adding 'apex/contrib/sparsity/permutation_search_kernels/call_permutation_sear
ch_kernels.py'
    adding 'apex/contrib/sparsity/permutation_search_kernels/exhaustive_search.py'
    adding
'apex/contrib/sparsity/permutation_search_kernels/permutation_utilities.py'
    adding 'apex/contrib/transducer/__init__.py'
    adding 'apex/contrib/transducer/transducer.py'
    adding 'apex/contrib/xentropy/__init__.py'
    adding 'apex/contrib/xentropy/softmax_xentropy.py'
    adding 'apex/fp16_utils/__init__.py'
    adding 'apex/fp16_utils/fp16_optimizer.py'
    adding 'apex/fp16_utils/fp16util.py'
    adding 'apex/fp16_utils/loss_scaler.py'
    adding 'apex/fused_dense/__init__.py'
    adding 'apex/fused_dense/fused_dense.py'
    adding 'apex/mlp/__init__.py'
    adding 'apex/mlp/mlp.py'
    adding 'apex/multi_tensor_apply/__init__.py'
    adding 'apex/multi_tensor_apply/multi_tensor_apply.py'
    adding 'apex/normalization/__init__.py'
    adding 'apex/normalization/fused_layer_norm.py'
    adding 'apex/optimizers/__init__.py'
    adding 'apex/optimizers/fused_adagrad.py'
    adding 'apex/optimizers/fused_adam.py'
    adding 'apex/optimizers/fused_lamb.py'
    adding 'apex/optimizers/fused_mixed_precision_lamb.py'
    adding 'apex/optimizers/fused_novograd.py'
    adding 'apex/optimizers/fused_sgd.py'

```

```

adding 'apex/parallel/LARC.py'
adding 'apex/parallel/__init__.py'
adding 'apex/parallel/distributed.py'
adding 'apex/parallel/multiproc.py'
adding 'apex/parallel/optimized_sync_batchnorm.py'
adding 'apex/parallel/optimized_sync_batchnorm_kernel.py'
adding 'apex/parallel/sync_batchnorm.py'
adding 'apex/parallel/sync_batchnorm_kernel.py'
adding 'apex/transformer/__init__.py'
adding 'apex/transformer/enums.py'
adding 'apex/transformer/log_util.py'
adding 'apex/transformer/microbatches.py'
adding 'apex/transformer/parallel_state.py'
adding 'apex/transformer/utils.py'
adding 'apex/transformer/_data/__init__.py'
adding 'apex/transformer/_data/_batch_sampler.py'
adding 'apex/transformer/amp/__init__.py'
adding 'apex/transformer/amp/grad_scaler.py'
adding 'apex/transformer/functional/__init__.py'
adding 'apex/transformer/functional/fused_softmax.py'
adding 'apex/transformer/layers/__init__.py'
adding 'apex/transformer/layers/layer_norm.py'
adding 'apex/transformer/pipeline_parallel/__init__.py'
adding 'apex/transformer/pipeline_parallel/_timers.py'
adding 'apex/transformer/pipeline_parallel/p2p_communication.py'
adding 'apex/transformer/pipeline_parallel/utils.py'
adding 'apex/transformer/pipeline_parallel/schedules/__init__.py'
adding 'apex/transformer/pipeline_parallel/schedules/common.py'
adding 'apex/transformer/pipeline_parallel/schedules/fwd_bwd_no_pipelining.py'
adding 'apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipelining_with_i
nterleaving.py'
adding 'apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipelining_withou
t_interleaving.py'
adding 'apex/transformer/tensor_parallel/__init__.py'
adding 'apex/transformer/tensor_parallel/cross_entropy.py'
adding 'apex/transformer/tensor_parallel/data.py'
adding 'apex/transformer/tensor_parallel/layers.py'
adding 'apex/transformer/tensor_parallel/mappings.py'
adding 'apex/transformer/tensor_parallel/memory.py'
adding 'apex/transformer/tensor_parallel/random.py'
adding 'apex/transformer/tensor_parallel/utils.py'
adding 'apex/transformer/testing/__init__.py'
adding 'apex/transformer/testing/arguments.py'
adding 'apex/transformer/testing/commons.py'
adding 'apex/transformer/testing/distributed_test_base.py'
adding 'apex/transformer/testing/global_vars.py'
adding 'apex/transformer/testing/standalone_bert.py'
adding 'apex/transformer/testing/standalone_gpt.py'

```



```

adding 'apex/transformer/testing/standalone_transformer_lm.py'
adding 'apex-0.1.dist-info/LICENSE'
adding 'apex-0.1.dist-info/METADATA'
adding 'apex-0.1.dist-info/WHEEL'
adding 'apex-0.1.dist-info/top_level.txt'
adding 'apex-0.1.dist-info/RECORD'
removing build/bdist.linux-x86_64/wheel
done
Created wheel for apex: filename=apex-0.1-py3-none-any.whl size=298114
sha256=2da4445c4b8d0aa9ae79dc8cccd27505ea9a7dcd7d9dc8dc400f6c6204e3e5ea
Stored in directory: /tmp/pip-ephem-wheel-cache-
pmttgsbe/wheels/6b/22/a2/1665526ee3c3061243260fd522525eca31398e04bfa5ad7e1f
Successfully built apex
Installing collected packages: apex
Value for scheme.platlib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.purelib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/apex
sysconfig: /usr/include/python3.7m/apex
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None

Value for scheme.platlib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.purelib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages

```

```

Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/UNKNOWN
sysconfig: /usr/include/python3.7m/UNKNOWN
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None
Successfully installed apex-0.1
Removed build tracker: '/tmp/pip-req-tracker-h_el88v4'
/content/ditto

```

```

[ ]: # some issue with colab
!pip install --upgrade "urllib3==1.25.4" awscli

```

```

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting urllib3==1.25.4
  Downloading urllib3-1.25.4-py2.py3-none-any.whl (125 kB)
    |                                     | 125 kB 5.3 MB/s
Collecting awscli
  Downloading awscli-1.25.46-py3-none-any.whl (3.9 MB)
    |                                     | 3.9 MB 47.2 MB/s
Collecting docutils<0.17,>=0.10
  Downloading docutils-0.16-py2.py3-none-any.whl (548 kB)
    |                                     | 548 kB 67.6 MB/s
Collecting colorama<0.4.5,>=0.2.5
  Downloading colorama-0.4.4-py2.py3-none-any.whl (16 kB)
Collecting s3transfer<0.7.0,>=0.6.0
  Downloading s3transfer-0.6.0-py3-none-any.whl (79 kB)
    |                                     | 79 kB 10.0 MB/s
Requirement already satisfied: PyYAML<5.5,>=3.10 in
/usr/local/lib/python3.7/dist-packages (from awscli) (3.13)
Collecting botocore==1.27.46
  Downloading botocore-1.27.46-py3-none-any.whl (9.0 MB)
    |                                     | 9.0 MB 44.0 MB/s
Collecting rsa<4.8,>=3.1.2
  Downloading rsa-4.7.2-py3-none-any.whl (34 kB)
Collecting jmespath<2.0.0,>=0.7.1

```

```
Downloading jmespath-1.0.1-py3-none-any.whl (20 kB)
Requirement already satisfied: python-dateutil<3.0.0,>=2.1 in
/usr/local/lib/python3.7/dist-packages (from botocore==1.27.46->awscli) (2.8.2)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.7/dist-
packages (from python-dateutil<3.0.0,>=2.1->botocore==1.27.46->awscli) (1.15.0)
Requirement already satisfied: pyasn1>=0.1.3 in /usr/local/lib/python3.7/dist-
packages (from rsa<4.8,>=3.1.2->awscli) (0.4.8)
Installing collected packages: urllib3, jmespath, botocore, s3transfer, rsa,
docutils, colorama, awscli
```

```
Attempting uninstall: urllib3
Found existing installation: urllib3 1.24.3
Uninstalling urllib3-1.24.3:
Successfully uninstalled urllib3-1.24.3
```

```
Attempting uninstall: rsa
Found existing installation: rsa 4.9
Uninstalling rsa-4.9:
Successfully uninstalled rsa-4.9
```

```
Attempting uninstall: docutils
Found existing installation: docutils 0.17.1
Uninstalling docutils-0.17.1:
Successfully uninstalled docutils-0.17.1
```

ERROR: pip's dependency resolver does not currently take into account all  
the packages that are installed. This behaviour is the source of the following  
dependency conflicts.

panel 0.12.1 requires tqdm>=4.48.0, but you have tqdm 4.41.0 which is  
incompatible.

fastai 2.7.7 requires torch<1.13,>=1.7, but you have torch 1.4.0 which is  
incompatible.

Successfully installed awscli-1.25.46 botocore-1.27.46 colorama-0.4.4  
docutils-0.16 jmespath-1.0.1 rsa-4.7.2 s3transfer-0.6.0 urllib3-1.25.4

```
[ ]: !pip install torch==1.7.1
```

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
```

```
Collecting torch==1.7.1
```

```
Downloading torch-1.7.1-cp37-cp37m-manylinux1_x86_64.whl (776.8 MB)
|                                     | 776.8 MB 17 kB/s
```

```
Requirement already satisfied: typing-extensions in
/usr/local/lib/python3.7/dist-packages (from torch==1.7.1) (4.1.1)
```

```
Requirement already satisfied: numpy in /usr/local/lib/python3.7/dist-packages
(from torch==1.7.1) (1.19.2)
```

```
Installing collected packages: torch
```

```
Attempting uninstall: torch
Found existing installation: torch 1.4.0
```

```

Uninstalling torch-1.4.0:
  Successfully uninstalled torch-1.4.0
ERROR: pip's dependency resolver does not currently take into account all
the packages that are installed. This behaviour is the source of the following
dependency conflicts.

torchvision 0.13.0+cu113 requires torch==1.12.0, but you have torch 1.7.1 which
is incompatible.

torchtext 0.13.0 requires torch==1.12.0, but you have torch 1.7.1 which is
incompatible.

torchaudio 0.12.0+cu113 requires torch==1.12.0, but you have torch 1.7.1 which
is incompatible.
Successfully installed torch-1.7.1

```

### 0.3 importa config.json

```
[ ]: !git clone https://github.com/pauloh48/IC.git
```

```

Cloning into 'IC'...
remote: Enumerating objects: 982, done.
remote: Counting objects: 100% (52/52), done.
remote: Compressing objects: 100% (48/48), done.
remote: Total 982 (delta 16), reused 7 (delta 0), pack-reused 930
Receiving objects: 100% (982/982), 251.89 MiB | 30.34 MiB/s, done.
Resolving deltas: 100% (600/600), done.
Checking out files: 100% (1273/1273), done.

```

### 0.4 remove e move

```
[ ]: !ls
```

```

apex          ditto.jpg    LICENSE     requirements.txt  run_all_wdc.py
blocking      ditto_light matcher.py  results_ditto    Snippext_public
configs.json  IC          output      run_all_er_magellan.py train_ditto.py
data          input       README.md   run_all_vary_size.py

```

```
[ ]: !rm configs.json
!mv -n /content/ditto/IC/datasEsErros/CMedium_30_3_5/configs.json /content/
↪ditto/
```

# 1 DA

## 1.1 DA STRUCTURED

### 1.1.1 Beer - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/Beer \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm distilbert \
  --da attr_del \
  --summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Beer.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/Beer \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm distilbert \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/
```

```
Downloading: 100% 442/442 [00:00<00:00, 373kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 349kB/s]
Downloading: 100% 268M/268M [00:03<00:00, 68.0MB/s]
step: 0, loss: 0.8741743564605713
epoch 1: dev_f1=0.2692307692307693, f1=0.28571428571428575,
best_f1=0.28571428571428575
step: 0, loss: 0.3771706223487854
epoch 2: dev_f1=0.2692307692307693, f1=0.27083333333333337,
best_f1=0.28571428571428575
step: 0, loss: 0.35589686036109924
epoch 3: dev_f1=0.2857142857142857, f1=0.0, best_f1=0.0
step: 0, loss: 0.3511686325073242
```

```

epoch 4: dev_f1=0.3333333333333333, f1=0.2222222222222222,
best_f1=0.2222222222222222
step: 0, loss: 0.2726292610168457
epoch 5: dev_f1=0.35135135135135137, f1=0.2631578947368421,
best_f1=0.2631578947368421
step: 0, loss: 0.2878105044364929
epoch 6: dev_f1=0.42857142857142855, f1=0.2857142857142857,
best_f1=0.2857142857142857
step: 0, loss: 0.2313859760761261
epoch 7: dev_f1=0.40625, f1=0.36363636363636365, best_f1=0.2857142857142857
step: 0, loss: 0.3241150677204132
epoch 8: dev_f1=0.4666666666666667, f1=0.13333333333333333,
best_f1=0.13333333333333333
step: 0, loss: 0.18947476148605347
epoch 9: dev_f1=0.42857142857142855, f1=0.125, best_f1=0.13333333333333333
step: 0, loss: 0.21668188273906708
epoch 10: dev_f1=0.4285714285714286, f1=0.2857142857142857,
best_f1=0.13333333333333333
step: 0, loss: 0.2083175778388977
epoch 11: dev_f1=0.48484848484848486, f1=0.22222222222222224,
best_f1=0.22222222222222224
step: 0, loss: 0.2005588561296463
epoch 12: dev_f1=0.5, f1=0.11764705882352941, best_f1=0.11764705882352941
step: 0, loss: 0.12714475393295288
epoch 13: dev_f1=0.4827586206896552, f1=0.11764705882352941,
best_f1=0.11764705882352941
step: 0, loss: 0.1421613246202469
epoch 14: dev_f1=0.5161290322580646, f1=0.11764705882352941,
best_f1=0.11764705882352941
step: 0, loss: 0.1826896220445633
epoch 15: dev_f1=0.5161290322580646, f1=0.11764705882352941,
best_f1=0.11764705882352941
mv: cannot stat '*_dev.pt': No such file or directory
Traceback (most recent call last):
  File "matcher.py", line 315, in <module>
    hp.lm, hp.use_gpu, hp.fp16)
  File "matcher.py", line 274, in load_model
    raise ModelNotFoundError(checkpoint)
ditto_light.exceptions.ModelNotFoundError: Model
checkpoints/Structured/Beer/model.pt was not found

```

### 1.1.2 DBLP-ACM - Running the matcher

```

[ ]: # Step 1:
      # train the wdc model with the save_model flag on
      # it will produce two files *_dev.pt and *_test.pt which are the

```

```

# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/DBLP-ACM \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm distilbert \
  --fp16 \
  --save_model \
  --da entry_swap \
  --da span_shuffle \
  --summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-ACM.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/DBLP-ACM \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm distilbert \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True

```

```

loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.8149300217628479
step: 10, loss: 0.4887552559375763
step: 20, loss: 0.5920388698577881
step: 30, loss: 0.40585941076278687
step: 40, loss: 0.1817753165960312
step: 50, loss: 0.12397924065589905
step: 60, loss: 0.10484205186367035
step: 70, loss: 0.097569040954113
step: 80, loss: 0.21009108424186707
step: 90, loss: 0.016763074323534966
step: 100, loss: 0.039560362696647644
step: 110, loss: 0.04549524560570717
step: 120, loss: 0.03419874981045723
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 130, loss: 0.038552042096853256
step: 140, loss: 0.05849090591073036
step: 150, loss: 0.0521862618625164
step: 160, loss: 0.22108280658721924
step: 170, loss: 0.0025654826313257217
step: 180, loss: 0.016393417492508888
step: 190, loss: 0.029518071562051773
step: 200, loss: 0.010128190740942955
step: 210, loss: 0.00955529510974884
step: 220, loss: 0.007362623233348131
step: 230, loss: 0.007180595304816961
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.

```



```

    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9763779527559054, f1=0.9727891156462585,
best_f1=0.9727891156462585
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015926355496048927
step: 10, loss: 0.0013487839605659246
step: 20, loss: 0.0043251048773527145
step: 30, loss: 0.0044152457267045975
step: 40, loss: 0.005140157416462898
step: 50, loss: 0.06531447172164917
step: 60, loss: 0.00723480386659503
step: 70, loss: 0.007304393220692873
step: 80, loss: 0.04297192394733429
step: 90, loss: 0.00245546898804605
step: 100, loss: 0.03388254716992378
step: 110, loss: 0.0018239641794934869
step: 120, loss: 0.024223260581493378
step: 130, loss: 0.004059189930558205
step: 140, loss: 0.00940192211419344
step: 150, loss: 0.004654275253415108
step: 160, loss: 0.022410903126001358
step: 170, loss: 0.09573562443256378
step: 180, loss: 0.0062592243775725365
step: 190, loss: 0.09843910485506058
step: 200, loss: 0.002482377691194415
step: 210, loss: 0.0349712148308754
step: 220, loss: 0.0024927714839577675
step: 230, loss: 0.03475864976644516
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.978675645342312, f1=0.9774266365688488,
best_f1=0.9774266365688488
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10704778879880905
step: 10, loss: 0.008386063389480114
step: 20, loss: 0.0033823116682469845
step: 30, loss: 0.013976500369608402
step: 40, loss: 0.025136740878224373
step: 50, loss: 0.06443969905376434

```

```

step: 60, loss: 0.005490707233548164
step: 70, loss: 0.0019167637219652534
step: 80, loss: 0.07024109363555908
step: 90, loss: 0.002588547533378005
step: 100, loss: 0.0008479873649775982
step: 110, loss: 0.03581807017326355
step: 120, loss: 0.004105409141629934
step: 130, loss: 0.0006444624159485102
step: 140, loss: 0.0006997095188125968
step: 150, loss: 0.001283789984881878
step: 160, loss: 0.0028065175283700228
step: 170, loss: 0.053906265646219254
step: 180, loss: 0.06490196287631989
step: 190, loss: 0.0017936421791091561
step: 200, loss: 0.03456595540046692
step: 210, loss: 0.0012604009825736284
step: 220, loss: 0.002307639690116048
step: 230, loss: 0.17202430963516235
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.9841628959276018, f1=0.9775280898876404,
best_f1=0.9775280898876404
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001402990659698844
step: 10, loss: 0.001427871291525662
step: 20, loss: 0.0026262816973030567
step: 30, loss: 0.0019466194789856672
step: 40, loss: 0.0036849265452474356
step: 50, loss: 0.05706196278333664
step: 60, loss: 0.010897970758378506
step: 70, loss: 0.007879738695919514
step: 80, loss: 0.057099997997283936
step: 90, loss: 0.0058690221048891544
step: 100, loss: 0.009234534576535225
step: 110, loss: 0.0022508446127176285
step: 120, loss: 0.048737701028585434
step: 130, loss: 0.002761267125606537
step: 140, loss: 0.0029579894617199898
step: 150, loss: 0.0024941496085375547
step: 160, loss: 0.0009320784010924399
step: 170, loss: 0.014770293608307838

```

```

step: 180, loss: 0.0894714817404747
step: 190, loss: 0.014774907380342484
step: 200, loss: 0.0013274719240143895
step: 210, loss: 0.00020143605070188642
step: 220, loss: 0.0005760747590102255
step: 230, loss: 0.003940131049603224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.9830890642615557, f1=0.9809203142536477,
best_f1=0.9775280898876404
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004711024521384388
step: 10, loss: 0.0025553342420607805
step: 20, loss: 0.0012314534978941083
step: 30, loss: 0.0141115328297019
step: 40, loss: 0.0007325423648580909
step: 50, loss: 0.0004988060682080686
step: 60, loss: 0.00659192306920886
step: 70, loss: 0.0004183224809821695
step: 80, loss: 0.0005683146300725639
step: 90, loss: 0.027614811435341835
step: 100, loss: 0.004082504194229841
step: 110, loss: 0.000348989648045972
step: 120, loss: 0.03765779361128807
step: 130, loss: 0.002720386255532503
step: 140, loss: 0.00036151870153844357
step: 150, loss: 0.0003240243240725249
step: 160, loss: 0.0009616721072234213
step: 170, loss: 0.02095225267112255
step: 180, loss: 0.0007074265740811825
step: 190, loss: 0.003539832541719079
step: 200, loss: 0.0217851921916008
step: 210, loss: 0.0019300914136692882
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 220, loss: 0.0005991609068587422
step: 230, loss: 0.00034350657369941473
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.9819413092550789, f1=0.9807037457434733,
best_f1=0.9775280898876404
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007407795637845993

```

```

step: 10, loss: 0.014616334810853004
step: 20, loss: 0.0004715705581475049
step: 30, loss: 0.00020071835024282336
step: 40, loss: 0.0005102345603518188
step: 50, loss: 0.0006489944644272327
step: 60, loss: 0.010417290963232517
step: 70, loss: 0.00048613507533445954
step: 80, loss: 0.00042396769276820123
step: 90, loss: 0.0009953547269105911
step: 100, loss: 0.00031810146174393594
step: 110, loss: 0.1695559173822403
step: 120, loss: 0.00021851898054592311
step: 130, loss: 0.002646172419190407
step: 140, loss: 0.18348337709903717
step: 150, loss: 0.0012999526225030422
step: 160, loss: 0.21864216029644012
step: 170, loss: 0.0005255054566077888
step: 180, loss: 0.008571183308959007
step: 190, loss: 0.0029208611231297255
step: 200, loss: 0.001885071280412376
step: 210, loss: 0.00607290118932724
step: 220, loss: 0.0010941404616460204
step: 230, loss: 0.00043950052349828184
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9794988610478361, f1=0.9748283752860413,
best_f1=0.9775280898876404
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007263513281941414
step: 10, loss: 0.00087897659977898
step: 20, loss: 0.0019298678962513804
step: 30, loss: 0.0005637943977490067
step: 40, loss: 0.05955186113715172
step: 50, loss: 0.00013506300456356257
step: 60, loss: 0.12185292690992355
step: 70, loss: 0.0007441457128152251
step: 80, loss: 0.0012758509255945683
step: 90, loss: 0.0022455391008406878
step: 100, loss: 0.01292648259550333
step: 110, loss: 0.002409161999821663
step: 120, loss: 0.001270002918317914
step: 130, loss: 0.006545319687575102
step: 140, loss: 0.00042995548574253917
step: 150, loss: 0.0003569632244762033
step: 160, loss: 0.00021089936490170658

```

```

step: 170, loss: 0.00014726698282174766
step: 180, loss: 0.10355336219072342
step: 190, loss: 0.0008492589113302529
step: 200, loss: 0.0025214198976755142
step: 210, loss: 0.00046692800242453814
step: 220, loss: 0.0008964447188191116
step: 230, loss: 0.024588869884610176
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9718785151856018, f1=0.9752252252252253,
best_f1=0.9775280898876404
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.029180625453591347
step: 10, loss: 0.01820901222527027
step: 20, loss: 0.0020782409701496363
step: 30, loss: 0.000723574950825423
step: 40, loss: 0.00038918235804885626
step: 50, loss: 0.0004844032519031316
step: 60, loss: 0.00027168745873495936
step: 70, loss: 0.0005666487850248814
step: 80, loss: 0.001437959959730506
step: 90, loss: 0.00018918211571872234
step: 100, loss: 0.00041767876246012747
step: 110, loss: 0.00028471110272221267
step: 120, loss: 0.0002395910123595968
step: 130, loss: 0.009325692430138588
step: 140, loss: 0.009384111501276493
step: 150, loss: 0.0003234184405300766
step: 160, loss: 0.0011924923164770007
step: 170, loss: 0.00030673586297780275
step: 180, loss: 0.003547398839145899
step: 190, loss: 0.0025378535501658916
step: 200, loss: 0.0033721146173775196
step: 210, loss: 0.0001780966849764809
step: 220, loss: 0.0034189624711871147
step: 230, loss: 0.0062889643013477325
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9830124575311437, f1=0.9797297297297298,
best_f1=0.9775280898876404
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007814778364263475

```

```

step: 10, loss: 0.00014883518451824784
step: 20, loss: 0.0007573359762318432
step: 30, loss: 0.00034308814792893827
step: 40, loss: 0.00031340427813120186
step: 50, loss: 0.000530891353264451
step: 60, loss: 0.00014529367035720497
step: 70, loss: 0.0006647386471740901
step: 80, loss: 0.011552141979336739
step: 90, loss: 0.001487608184106648
step: 100, loss: 0.0006458100979216397
step: 110, loss: 0.00034864217741414905
step: 120, loss: 0.024954557418823242
step: 130, loss: 0.00015143761993385851
step: 140, loss: 0.00011728621029760689
step: 150, loss: 9.187645628117025e-05
step: 160, loss: 0.00045439586392603815
step: 170, loss: 0.00013068370753899217
step: 180, loss: 0.000671949062962085
step: 190, loss: 0.00018581037875264883
step: 200, loss: 0.000481537077575922
step: 210, loss: 0.00016295518435072154
step: 220, loss: 0.0660267248749733
step: 230, loss: 0.00418467540293932
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9761634506242906, f1=0.9738933030646991,
best_f1=0.9775280898876404
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00027902130386792123
step: 10, loss: 0.0002546293253544718
step: 20, loss: 0.0022772387601435184
step: 30, loss: 0.0004798963200300932
step: 40, loss: 0.0001805144565878436
step: 50, loss: 0.0379922054708004
step: 60, loss: 0.022931329905986786
step: 70, loss: 0.000624916807282716
step: 80, loss: 0.0012545675272122025
step: 90, loss: 0.00022608083963859826
step: 100, loss: 0.00041103994590230286
step: 110, loss: 0.0001698080013738945
step: 120, loss: 0.006248195189982653
step: 130, loss: 0.00016598803631495684
step: 140, loss: 0.03022823855280876
step: 150, loss: 0.00012057575077051297
step: 160, loss: 0.0007281315047293901

```

```

step: 170, loss: 0.0011398348724469543
step: 180, loss: 0.00040655661723576486
step: 190, loss: 0.00017791423306334764
step: 200, loss: 7.047544931992888e-05
step: 210, loss: 0.00021087043569423258
step: 220, loss: 0.00024833937641233206
step: 230, loss: 0.0001036782850860618
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9808773903262092, f1=0.9774774774774775,
best_f1=0.9775280898876404
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000162877855473198
step: 10, loss: 0.00015955242270138115
step: 20, loss: 5.875428905710578e-05
step: 30, loss: 0.00042044659494422376
step: 40, loss: 0.007528937421739101
step: 50, loss: 0.002188602229580283
step: 60, loss: 0.0001382351911161095
step: 70, loss: 0.00020111750927753747
step: 80, loss: 0.00017389508138876408
step: 90, loss: 5.669975507771596e-05
step: 100, loss: 0.14288625121116638
step: 110, loss: 9.92012646747753e-05
step: 120, loss: 0.00010133824980584905
step: 130, loss: 6.309794844128191e-05
step: 140, loss: 6.730863242410123e-05
step: 150, loss: 0.0002706929517444223
step: 160, loss: 0.00010527704580454156
step: 170, loss: 0.02487887628376484
step: 180, loss: 0.00817484874278307
step: 190, loss: 0.002047909889370203
step: 200, loss: 0.00016667533782310784
step: 210, loss: 0.00012061371671734378
step: 220, loss: 0.00024846583255566657
step: 230, loss: 0.0019395905546844006
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9830890642615557, f1=0.9831649831649831,
best_f1=0.9775280898876404
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013063446385785937

```

```

step: 10, loss: 0.0005364222452044487
step: 20, loss: 6.266875425353646e-05
step: 30, loss: 0.00014396537153515965
step: 40, loss: 7.732795347692445e-05
step: 50, loss: 0.0001318114809691906
step: 60, loss: 0.004080838058143854
step: 70, loss: 0.00021341507090255618
step: 80, loss: 4.971025191480294e-05
step: 90, loss: 5.719220280298032e-05
step: 100, loss: 0.00014729455870110542
step: 110, loss: 0.00012671084550675005
step: 120, loss: 8.253542182501405e-05
step: 130, loss: 0.00021113452385179698
step: 140, loss: 0.00012235791655257344
step: 150, loss: 0.00011863566032843664
step: 160, loss: 0.026436205953359604
step: 170, loss: 9.532235708320513e-05
step: 180, loss: 0.00023646139015909284
step: 190, loss: 0.0001365435018669814
step: 200, loss: 0.0037846658378839493
step: 210, loss: 5.6608823797432706e-05
step: 220, loss: 0.0002294826990691945
step: 230, loss: 8.499303658027202e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.980963045912654, f1=0.9810055865921787,
best_f1=0.9775280898876404
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005317153409123421
step: 10, loss: 0.0001727024355204776
step: 20, loss: 9.20026795938611e-05
step: 30, loss: 0.027599314227700233
step: 40, loss: 9.273445175494999e-05
step: 50, loss: 0.00010244813893223181
step: 60, loss: 0.0001679752895142883
step: 70, loss: 0.0002228404046036303
step: 80, loss: 5.810436778119765e-05
step: 90, loss: 4.874175283475779e-05
step: 100, loss: 9.135340224020183e-05
step: 110, loss: 0.00010812567052198574
step: 120, loss: 0.0014869606820866466
step: 130, loss: 0.00012524781050160527
step: 140, loss: 5.465241338242777e-05
step: 150, loss: 0.020731007680296898
step: 160, loss: 5.371668885345571e-05

```



```

step: 170, loss: 5.7390861911699176e-05
step: 180, loss: 0.0002371031150687486
step: 190, loss: 4.267833355697803e-05
step: 200, loss: 5.096948007121682e-05
step: 210, loss: 6.678860518150032e-05
step: 220, loss: 5.756162136094645e-05
step: 230, loss: 4.250628626323305e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9819819819819819, f1=0.9809203142536477,
best_f1=0.9775280898876404
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017871921590995044
step: 10, loss: 4.177592199994251e-05
step: 20, loss: 0.02275185100734234
step: 30, loss: 0.00010825847130035982
step: 40, loss: 3.4580323699628934e-05
step: 50, loss: 5.403901741374284e-05
step: 60, loss: 5.60188309464138e-05
step: 70, loss: 0.00020024742116220295
step: 80, loss: 7.083707896526903e-05
step: 90, loss: 0.0002475949004292488
step: 100, loss: 0.0023947868030518293
step: 110, loss: 0.005459464155137539
step: 120, loss: 6.694138573948294e-05
step: 130, loss: 0.00022614069166593254
step: 140, loss: 0.00013574415061157197
step: 150, loss: 5.947049430687912e-05
step: 160, loss: 3.7571051507256925e-05
step: 170, loss: 5.5750620958860964e-05
step: 180, loss: 0.0001382952614221722
step: 190, loss: 0.0012144969077780843
step: 200, loss: 8.475200593238696e-05
step: 210, loss: 0.00025512499269098043
step: 220, loss: 8.728261309443042e-05
step: 230, loss: 3.6859302781522274e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9809203142536477, f1=0.9821029082774049,
best_f1=0.9775280898876404
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.5511431633494794e-05

```

```

step: 10, loss: 0.0005024885758757591
step: 20, loss: 7.323501631617546e-05
step: 30, loss: 0.00018611173436511308
step: 40, loss: 0.000146319463965483
step: 50, loss: 9.582907659932971e-05
step: 60, loss: 8.436720963800326e-05
step: 70, loss: 0.00039624894270673394
step: 80, loss: 6.361001578625292e-05
step: 90, loss: 3.478147118585184e-05
step: 100, loss: 8.367918781004846e-05
step: 110, loss: 7.417364395223558e-05
step: 120, loss: 5.120022615301423e-05
step: 130, loss: 6.119424506323412e-05
step: 140, loss: 5.541742575587705e-05
step: 150, loss: 7.268234912771732e-05
step: 160, loss: 4.8760335630504414e-05
step: 170, loss: 4.325706322561018e-05
step: 180, loss: 6.716707139275968e-05
step: 190, loss: 0.0035532298497855663
step: 200, loss: 4.920325955026783e-05
step: 210, loss: 0.004404084291309118
step: 220, loss: 5.2664232498500496e-05
step: 230, loss: 8.488849562127143e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9809203142536477, f1=0.9821029082774049,
best_f1=0.9775280898876404
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True

```

```

loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:10, 233.64it/s]
load_f1 = 0.983050847457627
real_f1 = 0.9830124575311437
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:16, 260.36it/s]

```

### 1.1.3 DBLP-GoogleScholar - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
    --task Structured/DBLP-GoogleScholar \
    --batch_size 32 \
    --max_len 256 \
    --lr 3e-5 \
    --n_epochs 15 \
    --finetuning \
    --lm distilbert \
    --fp16 \
    --save_model \
    --da entry_swap \
    --da span_shuffle \
    --summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-GoogleScholar.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
    --task Structured/DBLP-GoogleScholar \
    --input_path input/input_small.jsonl \
    --output_path output/output_small.jsonl \

```

```
--lm distilbert \  
--use_gpu \  
--fp16 \  
--checkpoint_path checkpoints/
```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```
enabled          : True  
opt_level        : 02  
cast_model_type  : torch.float16  
patch_torch_functions : False  
keep_batchnorm_fp32 : True  
master_weights   : True  
loss_scale       : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```
enabled          : True  
opt_level        : 02  
cast_model_type  : torch.float16  
patch_torch_functions : False  
keep_batchnorm_fp32 : True  
master_weights   : True  
loss_scale       : dynamic
```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0  
/usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`.

In PyTorch 1.1.0 and later, you should call them in the opposite order:

`optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

```
"https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",  
UserWarning)
```

```
step: 0, loss: 0.7994644641876221
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

```
step: 10, loss: 0.46882113814353943
```

```
step: 20, loss: 0.5023669004440308
```

```
step: 30, loss: 0.43852096796035767
```

```
step: 40, loss: 0.3518909215927124
```

step: 50, loss: 0.2948512136936188  
step: 60, loss: 0.2548460364341736  
step: 70, loss: 0.07825864851474762  
step: 80, loss: 0.24549753963947296  
step: 90, loss: 0.21410006284713745  
step: 100, loss: 0.29993292689323425  
step: 110, loss: 0.09243370592594147  
step: 120, loss: 0.07121257483959198  
step: 130, loss: 0.02435779757797718  
step: 140, loss: 0.23375175893306732  
step: 150, loss: 0.03062826208770275  
step: 160, loss: 0.10981424152851105  
step: 170, loss: 0.24090957641601562  
step: 180, loss: 0.1278885304927826  
step: 190, loss: 0.05255300551652908  
step: 200, loss: 0.10914001613855362  
step: 210, loss: 0.09734687954187393  
step: 220, loss: 0.082038015127182  
step: 230, loss: 0.14475010335445404  
step: 240, loss: 0.08959762006998062  
step: 250, loss: 0.07144138216972351  
step: 260, loss: 0.04006095230579376  
step: 270, loss: 0.024462847039103508  
step: 280, loss: 0.18683961033821106  
step: 290, loss: 0.10599177330732346  
step: 300, loss: 0.21167072653770447  
step: 310, loss: 0.13689494132995605  
step: 320, loss: 0.09285643696784973  
step: 330, loss: 0.17317086458206177  
step: 340, loss: 0.2326788604259491  
step: 350, loss: 0.0373457670211792  
step: 360, loss: 0.14072585105895996  
step: 370, loss: 0.1898503452539444  
step: 380, loss: 0.1564830094575882  
step: 390, loss: 0.0709085538983345  
step: 400, loss: 0.03646139055490494  
step: 410, loss: 0.04121106490492821  
step: 420, loss: 0.04420357570052147  
step: 430, loss: 0.06533564627170563  
step: 440, loss: 0.15917721390724182  
step: 450, loss: 0.03802633658051491  
step: 460, loss: 0.147110715508461  
step: 470, loss: 0.2558019161224365  
step: 480, loss: 0.2788561284542084  
step: 490, loss: 0.046869661659002304  
step: 500, loss: 0.02957092970609665  
step: 510, loss: 0.07120544463396072  
step: 520, loss: 0.14526332914829254

```

step: 530, loss: 0.09034769237041473
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9130434782608695, f1=0.9102682701202591,
best_f1=0.9102682701202591
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15015393495559692
step: 10, loss: 0.10749953240156174
step: 20, loss: 0.20834429562091827
step: 30, loss: 0.0794491246342659
step: 40, loss: 0.01782349869608879
step: 50, loss: 0.17950408160686493
step: 60, loss: 0.15780547261238098
step: 70, loss: 0.11459991335868835
step: 80, loss: 0.030110912397503853
step: 90, loss: 0.03423210233449936
step: 100, loss: 0.26941901445388794
step: 110, loss: 0.10054310411214828
step: 120, loss: 0.12447857111692429
step: 130, loss: 0.06309204548597336
step: 140, loss: 0.02237715944647789
step: 150, loss: 0.01667656935751438
step: 160, loss: 0.18336528539657593
step: 170, loss: 0.06947264820337296
step: 180, loss: 0.007140361703932285
step: 190, loss: 0.06299759447574615
step: 200, loss: 0.07865516096353531
step: 210, loss: 0.03419394791126251
step: 220, loss: 0.1724616140127182
step: 230, loss: 0.030442796647548676
step: 240, loss: 0.07206632196903229
step: 250, loss: 0.030712764710187912
step: 260, loss: 0.0028474254067987204
step: 270, loss: 0.2296433001756668
step: 280, loss: 0.3856066167354584
step: 290, loss: 0.12196892499923706
step: 300, loss: 0.028740989044308662
step: 310, loss: 0.0895855724811554
step: 320, loss: 0.07093305140733719
step: 330, loss: 0.04406075179576874
step: 340, loss: 0.009128594771027565

```

```

step: 350, loss: 0.03763722628355026
step: 360, loss: 0.029940858483314514
step: 370, loss: 0.009331979788839817
step: 380, loss: 0.06440871208906174
step: 390, loss: 0.09198514372110367
step: 400, loss: 0.06858761608600616
step: 410, loss: 0.0012754093622788787
step: 420, loss: 0.10710883140563965
step: 430, loss: 0.019536329433321953
step: 440, loss: 0.025082362815737724
step: 450, loss: 0.01221170462667942
step: 460, loss: 0.23220638930797577
step: 470, loss: 0.03462821617722511
step: 480, loss: 0.20648787915706635
step: 490, loss: 0.03482438251376152
step: 500, loss: 0.013411538675427437
step: 510, loss: 0.05381202697753906
step: 520, loss: 0.0029766480438411236
step: 530, loss: 0.12302812933921814
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9259606373008434, f1=0.9252336448598131,
best_f1=0.9252336448598131
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04311187192797661
step: 10, loss: 0.05228913575410843
step: 20, loss: 0.10696838051080704
step: 30, loss: 0.12872879207134247
step: 40, loss: 0.006050224415957928
step: 50, loss: 0.011190136894583702
step: 60, loss: 0.0076638213358819485
step: 70, loss: 0.0192171111702919
step: 80, loss: 0.01615169644355774
step: 90, loss: 0.018023254349827766
step: 100, loss: 0.012945153750479221
step: 110, loss: 0.01616726629436016
step: 120, loss: 0.02579023502767086
step: 130, loss: 0.016740625724196434
step: 140, loss: 0.025583554059267044
step: 150, loss: 0.014252953231334686

```

```

step: 160, loss: 0.00984116829931736
step: 170, loss: 0.011026880703866482
step: 180, loss: 0.010619795881211758
step: 190, loss: 0.0023850572761148214
step: 200, loss: 0.03674659505486488
step: 210, loss: 0.03358830139040947
step: 220, loss: 0.09770182520151138
step: 230, loss: 0.02016299217939377
step: 240, loss: 0.009408973157405853
step: 250, loss: 0.006898229010403156
step: 260, loss: 0.002083888975903392
step: 270, loss: 0.002867551753297448
step: 280, loss: 0.018801946192979813
step: 290, loss: 0.04196618124842644
step: 300, loss: 0.2501172125339508
step: 310, loss: 0.11007489264011383
step: 320, loss: 0.06590909510850906
step: 330, loss: 0.0015908109489828348
step: 340, loss: 0.0046508884988725185
step: 350, loss: 0.03342008963227272
step: 360, loss: 0.011739675886929035
step: 370, loss: 0.011832233518362045
step: 380, loss: 0.010758268646895885
step: 390, loss: 0.007103848736733198
step: 400, loss: 0.013178437948226929
step: 410, loss: 0.007422268856316805
step: 420, loss: 0.01410723477602005
step: 430, loss: 0.004965873435139656
step: 440, loss: 0.011386319063603878
step: 450, loss: 0.08401766419410706
step: 460, loss: 0.06706147640943527
step: 470, loss: 0.01128815021365881
step: 480, loss: 0.034549858421087265
step: 490, loss: 0.022781342267990112
step: 500, loss: 0.11894157528877258
step: 510, loss: 0.017703043296933174
step: 520, loss: 0.018339136615395546
step: 530, loss: 0.02631657011806965
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.9190968955785513, f1=0.9253034547152195,
best_f1=0.9252336448598131
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004607075359672308
step: 10, loss: 0.0059150345623493195

```



step: 20, loss: 0.0781816840171814  
step: 30, loss: 0.0011486017610877752  
step: 40, loss: 0.0029502774123102427  
step: 50, loss: 0.059246741235256195  
step: 60, loss: 0.0019056854071095586  
step: 70, loss: 0.001437878469005227  
step: 80, loss: 0.010503990575671196  
step: 90, loss: 0.011834817938506603  
step: 100, loss: 0.1105024591088295  
step: 110, loss: 0.004400791600346565  
step: 120, loss: 0.003946947865188122  
step: 130, loss: 0.008521811105310917  
step: 140, loss: 0.0427897572517395  
step: 150, loss: 0.00160409661475569  
step: 160, loss: 0.013822803273797035  
step: 170, loss: 0.005272410809993744  
step: 180, loss: 0.0028227386064827442  
step: 190, loss: 0.004718622658401728  
step: 200, loss: 0.0078971516340971  
step: 210, loss: 0.0017422351520508528  
step: 220, loss: 0.097127266228199  
step: 230, loss: 0.1983775645494461  
step: 240, loss: 0.01734563708305359  
step: 250, loss: 0.009836710058152676  
step: 260, loss: 0.08441346883773804  
step: 270, loss: 0.015001429244875908  
step: 280, loss: 0.004810620564967394  
step: 290, loss: 0.14198440313339233  
step: 300, loss: 0.006907665636390448  
step: 310, loss: 0.008676773868501186  
step: 320, loss: 0.0461452379822731  
step: 330, loss: 0.14702185988426208  
step: 340, loss: 0.016746509820222855  
step: 350, loss: 0.0008854255429469049  
step: 360, loss: 0.0018980238819494843  
step: 370, loss: 0.04122540354728699  
step: 380, loss: 0.001568949781358242  
step: 390, loss: 0.10751451551914215  
step: 400, loss: 0.04153188690543175  
step: 410, loss: 0.03722713515162468  
step: 420, loss: 0.0051127406768500805  
step: 430, loss: 0.008635486476123333  
step: 440, loss: 0.04008832946419716  
step: 450, loss: 0.014306796714663506  
step: 460, loss: 0.005659597460180521  
step: 470, loss: 0.003683279501274228  
step: 480, loss: 0.0016517817275598645  
step: 490, loss: 0.047644250094890594

```

step: 500, loss: 0.07479605078697205
step: 510, loss: 0.043696481734514236
step: 520, loss: 0.020165864378213882
step: 530, loss: 0.01122539583593607
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.9156853509185116, f1=0.9213483146067416,
best_f1=0.9252336448598131
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003492509014904499
step: 10, loss: 0.002447455655783415
step: 20, loss: 0.0097062261775136
step: 30, loss: 0.015142634510993958
step: 40, loss: 0.024471374228596687
step: 50, loss: 0.04362509027123451
step: 60, loss: 0.01396605558693409
step: 70, loss: 0.0005869847955182195
step: 80, loss: 0.0006463199388235807
step: 90, loss: 0.09283822774887085
step: 100, loss: 0.0005913391942158341
step: 110, loss: 0.016448577865958214
step: 120, loss: 0.3325546383857727
step: 130, loss: 0.005881778430193663
step: 140, loss: 0.010797960683703423
step: 150, loss: 0.01006791740655899
step: 160, loss: 0.04579104855656624
step: 170, loss: 0.004639674909412861
step: 180, loss: 0.0019213773775845766
step: 190, loss: 0.008782656863331795
step: 200, loss: 0.0010905893286690116
step: 210, loss: 0.001279020099900663
step: 220, loss: 0.00034559215418994427
step: 230, loss: 0.003736665239557624
step: 240, loss: 0.004112352151423693
step: 250, loss: 0.001642104471102357
step: 260, loss: 0.0005875161732546985
step: 270, loss: 0.06831151992082596
step: 280, loss: 0.07931055873632431
step: 290, loss: 0.0047719539143145084
step: 300, loss: 0.009635630995035172
step: 310, loss: 0.00881210993975401
step: 320, loss: 0.006215271074324846
step: 330, loss: 0.01637732796370983
step: 340, loss: 0.00839919876307249
step: 350, loss: 0.00966768991202116

```

```

step: 360, loss: 0.004302388988435268
step: 370, loss: 0.007508736569434404
step: 380, loss: 0.0599699430167675
step: 390, loss: 0.009878851473331451
step: 400, loss: 0.011673703789710999
step: 410, loss: 0.0007328451029025018
step: 420, loss: 0.05592738091945648
step: 430, loss: 0.007119120564311743
step: 440, loss: 0.0010504024103283882
step: 450, loss: 0.019267233088612556
step: 460, loss: 0.0004727386694867164
step: 470, loss: 0.0026291687972843647
step: 480, loss: 0.002405231585726142
step: 490, loss: 0.00895379763096571
step: 500, loss: 0.04978936165571213
step: 510, loss: 0.01304540690034628
step: 520, loss: 0.0012332655023783445
step: 530, loss: 0.00481836823746562
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.9162011173184358, f1=0.9242843951985227,
best_f1=0.9252336448598131
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0039835902862250805
step: 10, loss: 0.0017689428059384227
step: 20, loss: 0.0013265633024275303
step: 30, loss: 0.016867903992533684
step: 40, loss: 0.0012340539833530784
step: 50, loss: 0.0013436774024739861
step: 60, loss: 0.0006975043215788901
step: 70, loss: 0.06646489351987839
step: 80, loss: 0.000285860151052475
step: 90, loss: 0.008672797121107578
step: 100, loss: 0.06728223711252213
step: 110, loss: 0.0038561816327273846
step: 120, loss: 0.0013002518098801374
step: 130, loss: 0.00048482557758688927
step: 140, loss: 0.0009579831967130303
step: 150, loss: 0.15567578375339508
step: 160, loss: 0.000737042457330972
step: 170, loss: 0.0002937997633125633
step: 180, loss: 0.003832723945379257
step: 190, loss: 0.0016165025299414992
step: 200, loss: 0.0004532553721219301
step: 210, loss: 0.00026833752053789794

```

```

step: 220, loss: 0.009195693768560886
step: 230, loss: 0.00022903767239768058
step: 240, loss: 0.010560017079114914
step: 250, loss: 0.0002652388939168304
step: 260, loss: 0.002342954510822892
step: 270, loss: 0.0015538946026936173
step: 280, loss: 0.004237199202179909
step: 290, loss: 0.001003930578008294
step: 300, loss: 0.008616170845925808
step: 310, loss: 0.0004557094071060419
step: 320, loss: 0.05637756362557411
step: 330, loss: 0.0014910276513546705
step: 340, loss: 0.019567497074604034
step: 350, loss: 0.08719746768474579
step: 360, loss: 0.0003241216472815722
step: 370, loss: 0.027094589546322823
step: 380, loss: 0.031008288264274597
step: 390, loss: 0.1365300565958023
step: 400, loss: 0.0005716991145163774
step: 410, loss: 0.0012529522646218538
step: 420, loss: 0.01558581180870533
step: 430, loss: 0.0007299157441593707
step: 440, loss: 0.00950026698410511
step: 450, loss: 0.01802833378314972
step: 460, loss: 0.08238893747329712
step: 470, loss: 0.0036281466018408537
step: 480, loss: 0.0061238170601427555
step: 490, loss: 0.0033115746919065714
step: 500, loss: 0.006827642675489187
step: 510, loss: 0.0006314473575912416
step: 520, loss: 0.0258912555873394
step: 530, loss: 0.019345510751008987
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9164700330344502, f1=0.915014164305949,
best_f1=0.9252336448598131
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007821359322406352
step: 10, loss: 0.034575752913951874
step: 20, loss: 0.024647358804941177
step: 30, loss: 0.0003961576730944216
step: 40, loss: 0.0006138482131063938
step: 50, loss: 0.08753988146781921
step: 60, loss: 0.0001871799468062818
step: 70, loss: 0.02519674226641655

```

```
step: 80, loss: 0.04855607822537422
step: 90, loss: 0.00033705533132888377
step: 100, loss: 0.0022359532304108143
step: 110, loss: 0.019727973267436028
step: 120, loss: 0.00033236207673326135
step: 130, loss: 0.00017589589697308838
step: 140, loss: 0.00041403714567422867
step: 150, loss: 8.444830746157095e-05
step: 160, loss: 0.00017300600302405655
step: 170, loss: 0.00016497827891726047
step: 180, loss: 0.023805901408195496
step: 190, loss: 0.0001922452065628022
step: 200, loss: 0.08892415463924408
step: 210, loss: 0.006281980779021978
step: 220, loss: 0.0003830378409475088
step: 230, loss: 0.0003330419131089002
step: 240, loss: 0.008888998068869114
step: 250, loss: 0.0011384427780285478
step: 260, loss: 0.007791834883391857
step: 270, loss: 0.000952667323872447
step: 280, loss: 0.05278070643544197
step: 290, loss: 0.0044815558940172195
step: 300, loss: 0.007188244257122278
step: 310, loss: 0.00014497913070954382
step: 320, loss: 0.14483346045017242
step: 330, loss: 0.00019100337522104383
step: 340, loss: 0.01766090653836727
step: 350, loss: 0.0009733364568091929
step: 360, loss: 0.04937832057476044
step: 370, loss: 0.0008219768642447889
step: 380, loss: 0.002048675436526537
step: 390, loss: 0.00012931633682455868
step: 400, loss: 0.0013771580997854471
step: 410, loss: 0.0016588250873610377
step: 420, loss: 0.0003588031977415085
step: 430, loss: 7.027459650998935e-05
step: 440, loss: 0.0002342046791454777
step: 450, loss: 0.03234197199344635
step: 460, loss: 0.005455052014440298
step: 470, loss: 0.0009703607065603137
step: 480, loss: 0.0010350579395890236
step: 490, loss: 0.00038996993680484593
step: 500, loss: 4.7091645683394745e-05
step: 510, loss: 0.0004084007814526558
step: 520, loss: 0.011833030730485916
step: 530, loss: 0.00023389408306684345
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9195294117647058, f1=0.9209783631232361,
best_f1=0.9252336448598131
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03992145508527756
step: 10, loss: 0.0023513808846473694
step: 20, loss: 0.011781374923884869
step: 30, loss: 0.02030899003148079
step: 40, loss: 0.014101602137088776
step: 50, loss: 0.0005214769626036286
step: 60, loss: 0.00010578870569588616
step: 70, loss: 0.012630264274775982
step: 80, loss: 0.00023247908393386751
step: 90, loss: 0.00010612786718411371
step: 100, loss: 0.0032640104182064533
step: 110, loss: 0.0016844678902998567
step: 120, loss: 0.018409769982099533
step: 130, loss: 0.00039564547478221357
step: 140, loss: 3.889710933435708e-05
step: 150, loss: 0.0003341183764860034
step: 160, loss: 0.002927496563643217
step: 170, loss: 0.00130532740149647
step: 180, loss: 0.0006304421694949269
step: 190, loss: 0.00010650878539308906
step: 200, loss: 8.960279228631407e-05
step: 210, loss: 0.008735834620893002
step: 220, loss: 0.00017889468290377408
step: 230, loss: 0.00012815895024687052
step: 240, loss: 5.7188666687579826e-05
step: 250, loss: 0.0001062471346813254
step: 260, loss: 0.04635138064622879
step: 270, loss: 3.696481871884316e-05
step: 280, loss: 0.024388404563069344
step: 290, loss: 9.639560448704287e-05
step: 300, loss: 0.00021940180158708245
step: 310, loss: 0.036040958017110825
step: 320, loss: 0.00011100027768407017
step: 330, loss: 0.016912231221795082
step: 340, loss: 0.0003487500362098217
step: 350, loss: 0.004442051984369755
step: 360, loss: 0.0010334389517083764
step: 370, loss: 0.00019882649939972907
step: 380, loss: 0.014425056055188179
step: 390, loss: 0.00021107961947564036
step: 400, loss: 0.07071948796510696
step: 410, loss: 0.00013779630535282195
```

```

step: 420, loss: 0.0005461505497805774
step: 430, loss: 0.0026324999053031206
step: 440, loss: 0.005611399654299021
step: 450, loss: 0.00033693533623591065
step: 460, loss: 0.0005250314134173095
step: 470, loss: 0.00033563069882802665
step: 480, loss: 0.0024412504862993956
step: 490, loss: 0.00031015908461995423
step: 500, loss: 7.142753020161763e-05
step: 510, loss: 8.659311424707994e-05
step: 520, loss: 9.699531801743433e-05
step: 530, loss: 0.0018151712138205767
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9096209912536443, f1=0.9134010643444606,
best_f1=0.9252336448598131
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002619385486468673
step: 10, loss: 0.00045952899381518364
step: 20, loss: 0.00014500506222248077
step: 30, loss: 0.0006920424057170749
step: 40, loss: 0.034521233290433884
step: 50, loss: 0.00016020632756408304
step: 60, loss: 7.893201836850494e-05
step: 70, loss: 0.11595707386732101
step: 80, loss: 0.0037581694778054953
step: 90, loss: 0.00012332423648331314
step: 100, loss: 0.00015737964713480324
step: 110, loss: 0.0001515617623226717
step: 120, loss: 7.531574374297634e-05
step: 130, loss: 0.0007860012701712549
step: 140, loss: 0.05795967951416969
step: 150, loss: 7.552224269602448e-05
step: 160, loss: 0.002900925697758794
step: 170, loss: 0.00012747578148264438
step: 180, loss: 5.46653573110234e-05
step: 190, loss: 0.001996114384382963
step: 200, loss: 0.005380666349083185
step: 210, loss: 4.817378066945821e-05
step: 220, loss: 0.0006147746462374926
step: 230, loss: 5.7422024838160723e-05
step: 240, loss: 0.00033741656807251275
step: 250, loss: 7.952158921398222e-05
step: 260, loss: 0.00011962384451180696
step: 270, loss: 0.00033298583002761006

```

```

step: 280, loss: 6.059974475647323e-05
step: 290, loss: 4.914098099106923e-05
step: 300, loss: 7.273288065334782e-05
step: 310, loss: 4.0656559576746076e-05
step: 320, loss: 0.0002682497724890709
step: 330, loss: 0.00014053205086383969
step: 340, loss: 5.351808067644015e-05
step: 350, loss: 0.000768367201089859
step: 360, loss: 0.03208918496966362
step: 370, loss: 0.000168236656463705
step: 380, loss: 9.185713861370459e-05
step: 390, loss: 4.104306935914792e-05
step: 400, loss: 0.015581196174025536
step: 410, loss: 0.007997626438736916
step: 420, loss: 0.004410028923302889
step: 430, loss: 4.1530915041221306e-05
step: 440, loss: 0.009065698832273483
step: 450, loss: 0.0020496079232543707
step: 460, loss: 0.00634900014847517
step: 470, loss: 0.002043417189270258
step: 480, loss: 0.0012867363402619958
step: 490, loss: 6.010647848597728e-05
step: 500, loss: 0.010873522609472275
step: 510, loss: 0.0024957575369626284
step: 520, loss: 0.00016837920702528208
step: 530, loss: 9.251248411601409e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9202279202279201, f1=0.9214953271028037,
best_f1=0.9252336448598131
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.330699397949502e-05
step: 10, loss: 2.304028021171689e-05
step: 20, loss: 8.14335944596678e-05
step: 30, loss: 8.79651925060898e-05
step: 40, loss: 0.0005152174271643162
step: 50, loss: 0.0001420663029421121
step: 60, loss: 0.00032729972735978663
step: 70, loss: 0.0009765870636329055
step: 80, loss: 8.857519424054772e-05
step: 90, loss: 0.00022408788208849728
step: 100, loss: 6.540303729707375e-05
step: 110, loss: 0.0011922039557248354
step: 120, loss: 9.75525617832318e-05
step: 130, loss: 5.1702427299460396e-05

```



```

step: 140, loss: 0.001755217555910349
step: 150, loss: 3.859025673591532e-05
step: 160, loss: 4.531287413556129e-05
step: 170, loss: 0.0001226775348186493
step: 180, loss: 0.004706188570708036
step: 190, loss: 4.9776579544413835e-05
step: 200, loss: 0.0026105400174856186
step: 210, loss: 0.00011956864182138816
step: 220, loss: 0.0001000589836621657
step: 230, loss: 8.085645822575316e-05
step: 240, loss: 7.551161252195016e-05
step: 250, loss: 5.018958472646773e-05
step: 260, loss: 0.01629560999572277
step: 270, loss: 8.287827222375199e-05
step: 280, loss: 9.046029299497604e-05
step: 290, loss: 0.00022378198627848178
step: 300, loss: 0.0009684719261713326
step: 310, loss: 0.0013191996840760112
step: 320, loss: 9.13557450985536e-05
step: 330, loss: 0.001579187111929059
step: 340, loss: 0.0001707528717815876
step: 350, loss: 0.00015850561612751335
step: 360, loss: 0.0007835825090296566
step: 370, loss: 7.672196079511195e-05
step: 380, loss: 6.699597724946216e-05
step: 390, loss: 0.0004562629328574985
step: 400, loss: 0.0004980754456482828
step: 410, loss: 8.453496411675587e-05
step: 420, loss: 0.002450529020279646
step: 430, loss: 0.000208867626497522
step: 440, loss: 6.333497731247917e-05
step: 450, loss: 0.000749860773794353
step: 460, loss: 0.0002971018257085234
step: 470, loss: 4.207813253742643e-05
step: 480, loss: 0.0005183246685191989
step: 490, loss: 0.0639382004737854
step: 500, loss: 0.003296047216281295
step: 510, loss: 0.013786576688289642
step: 520, loss: 0.016119681298732758
step: 530, loss: 3.0020853955647908e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9241639189825719, f1=0.9295112781954888,
best_f1=0.9252336448598131
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 6.878667045384645e-05  
step: 10, loss: 0.0027479189448058605  
step: 20, loss: 2.422471698082518e-05  
step: 30, loss: 0.027230365201830864  
step: 40, loss: 0.0021062397863715887  
step: 50, loss: 0.0003238916106056422  
step: 60, loss: 9.556351869832724e-05  
step: 70, loss: 0.0005892403423786163  
step: 80, loss: 0.02110760286450386  
step: 90, loss: 0.00014005009143147618  
step: 100, loss: 7.219198596430942e-05  
step: 110, loss: 0.00010650521289790049  
step: 120, loss: 0.00012637287727557123  
step: 130, loss: 4.230541890137829e-05  
step: 140, loss: 5.2730993047589436e-05  
step: 150, loss: 7.944118260638788e-05  
step: 160, loss: 0.0025240846443921328  
step: 170, loss: 6.644934182986617e-05  
step: 180, loss: 4.3816860852530226e-05  
step: 190, loss: 0.009964422322809696  
step: 200, loss: 0.0011187250493094325  
step: 210, loss: 0.00010455261508468539  
step: 220, loss: 0.0001825051731429994  
step: 230, loss: 0.0019786739721894264  
step: 240, loss: 0.00017182859301101416  
step: 250, loss: 3.512758485157974e-05  
step: 260, loss: 0.00036109157372266054  
step: 270, loss: 0.00036833793274126947  
step: 280, loss: 0.0015715771587565541  
step: 290, loss: 0.00045609992230311036  
step: 300, loss: 0.016680749133229256  
step: 310, loss: 4.4568292651092634e-05  
step: 320, loss: 0.0217348150908947  
step: 330, loss: 0.0007312659290619195  
step: 340, loss: 6.30310969427228e-05  
step: 350, loss: 0.0007191275362856686  
step: 360, loss: 5.640329254674725e-05  
step: 370, loss: 5.672175029758364e-05  
step: 380, loss: 2.1710526198148727e-05  
step: 390, loss: 0.0021497374400496483  
step: 400, loss: 2.8560369173646905e-05  
step: 410, loss: 0.006944688968360424  
step: 420, loss: 8.750155393499881e-05  
step: 430, loss: 0.024717940017580986  
step: 440, loss: 0.00444381358101964  
step: 450, loss: 0.0001614471257198602  
step: 460, loss: 0.03894974663853645  
step: 470, loss: 0.0009440772701054811

```

step: 480, loss: 0.0010334409307688475
step: 490, loss: 0.0014716307632625103
step: 500, loss: 0.00012683992099482566
step: 510, loss: 0.0002032167394645512
step: 520, loss: 0.001109772827476263
step: 530, loss: 4.5948258048156276e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9185938945420906, f1=0.9275893675527039,
best_f1=0.9252336448598131
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.095862722257152e-05
step: 10, loss: 0.000569486350286752
step: 20, loss: 7.132943574106321e-05
step: 30, loss: 0.00021439572446979582
step: 40, loss: 0.00014842386008240283
step: 50, loss: 0.00018481550796423107
step: 60, loss: 0.0008176557021215558
step: 70, loss: 0.0004068883426953107
step: 80, loss: 0.0014974439982324839
step: 90, loss: 0.00019388037617318332
step: 100, loss: 0.008801248855888844
step: 110, loss: 4.3158299376955256e-05
step: 120, loss: 0.00013162176765035838
step: 130, loss: 0.001003566081635654
step: 140, loss: 4.960632213624194e-05
step: 150, loss: 0.005010439082980156
step: 160, loss: 0.000595640332903713
step: 170, loss: 1.6964670066954568e-05
step: 180, loss: 0.0012640804052352905
step: 190, loss: 9.383098222315311e-05
step: 200, loss: 9.159617911791429e-05
step: 210, loss: 0.001922281226143241
step: 220, loss: 0.00014505453873425722
step: 230, loss: 0.0006693948525935411
step: 240, loss: 0.0045493002980947495
step: 250, loss: 3.581683267839253e-05
step: 260, loss: 4.458178227650933e-05
step: 270, loss: 3.877720519085415e-05
step: 280, loss: 4.3250667658867314e-05
step: 290, loss: 0.0024240801576524973
step: 300, loss: 4.96595494041685e-05
step: 310, loss: 4.800092938239686e-05
step: 320, loss: 5.125831739860587e-05
step: 330, loss: 5.404548574006185e-05

```

```

step: 340, loss: 7.5122581620235e-05
step: 350, loss: 0.008290107361972332
step: 360, loss: 0.0001271888759220019
step: 370, loss: 4.0844231989467517e-05
step: 380, loss: 3.4382468584226444e-05
step: 390, loss: 2.4649451006553136e-05
step: 400, loss: 5.810840229969472e-05
step: 410, loss: 0.0009646436083130538
step: 420, loss: 0.001407602452673018
step: 430, loss: 9.478531137574464e-05
step: 440, loss: 8.262287155957893e-05
step: 450, loss: 0.0011866649147123098
step: 460, loss: 0.0005165633629076183
step: 470, loss: 5.484901339514181e-05
step: 480, loss: 4.28936182288453e-05
step: 490, loss: 0.015226666815578938
step: 500, loss: 0.003005025442689657
step: 510, loss: 0.0009451006771996617
step: 520, loss: 0.02339186891913414
step: 530, loss: 0.004693529102951288
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9183285849952516, f1=0.9187913125590179,
best_f1=0.9252336448598131
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.4798318918328732e-05
step: 10, loss: 0.0003220342332497239
step: 20, loss: 0.01537275779992342
step: 30, loss: 0.0015763579867780209
step: 40, loss: 4.08412488468457e-05
step: 50, loss: 8.097320096567273e-05
step: 60, loss: 3.3350799640174955e-05
step: 70, loss: 0.0001583004486747086
step: 80, loss: 3.764400753425434e-05
step: 90, loss: 0.0018295533955097198
step: 100, loss: 2.9772856578347273e-05
step: 110, loss: 4.438198448042385e-05
step: 120, loss: 7.658148388145491e-05
step: 130, loss: 0.011183828115463257
step: 140, loss: 2.162845339626074e-05
step: 150, loss: 0.0005236199358478189
step: 160, loss: 3.9781782106729224e-05
step: 170, loss: 0.0010064791422337294
step: 180, loss: 0.0004491849977057427
step: 190, loss: 5.0645718147279695e-05

```

```

step: 200, loss: 0.0008174823597073555
step: 210, loss: 0.0012849774211645126
step: 220, loss: 0.0005873769987374544
step: 230, loss: 0.0002030720643233508
step: 240, loss: 5.3378960728878155e-05
step: 250, loss: 0.05234546959400177
step: 260, loss: 0.001465525128878653
step: 270, loss: 2.2295313101494685e-05
step: 280, loss: 1.4971671589592006e-05
step: 290, loss: 2.6586099920677952e-05
step: 300, loss: 0.00017437127826269716
step: 310, loss: 0.00015338865341618657
step: 320, loss: 0.004326160531491041
step: 330, loss: 2.4347740691155195e-05
step: 340, loss: 3.686674972414039e-05
step: 350, loss: 0.03404462710022926
step: 360, loss: 9.214469173457474e-05
step: 370, loss: 1.804837302188389e-05
step: 380, loss: 6.872189260320738e-05
step: 390, loss: 3.3553933462826535e-05
step: 400, loss: 5.6527278502471745e-05
step: 410, loss: 2.547630538174417e-05
step: 420, loss: 2.736507485678885e-05
step: 430, loss: 0.00016063191287685186
step: 440, loss: 3.9968403143575415e-05
step: 450, loss: 4.5960092393215746e-05
step: 460, loss: 1.7311051124124788e-05
step: 470, loss: 0.004054140765219927
step: 480, loss: 0.0002276466111652553
step: 490, loss: 1.837653326219879e-05
step: 500, loss: 8.01813876023516e-05
step: 510, loss: 8.778263872954994e-05
step: 520, loss: 3.226672561140731e-05
step: 530, loss: 2.5454115530010313e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.919463087248322, f1=0.9183770883054891,
best_f1=0.9252336448598131
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.518214205338154e-05
step: 10, loss: 0.0003581464698072523
step: 20, loss: 1.999680716835428e-05
step: 30, loss: 0.00018243957310914993
step: 40, loss: 9.468811913393438e-05
step: 50, loss: 0.00024440104607492685

```

step: 60, loss: 2.4373912310693413e-05  
step: 70, loss: 2.21014615817694e-05  
step: 80, loss: 5.276228330330923e-05  
step: 90, loss: 1.6461728591821156e-05  
step: 100, loss: 2.921660961874295e-05  
step: 110, loss: 2.4059805582510307e-05  
step: 120, loss: 2.3458580471924506e-05  
step: 130, loss: 0.002984524006024003  
step: 140, loss: 4.074804382980801e-05  
step: 150, loss: 2.9399026971077546e-05  
step: 160, loss: 0.00010447816748637706  
step: 170, loss: 0.0005175807746127248  
step: 180, loss: 2.4507824491593055e-05  
step: 190, loss: 0.0006762270350009203  
step: 200, loss: 2.313734876224771e-05  
step: 210, loss: 3.925461714970879e-05  
step: 220, loss: 0.00011421708768466488  
step: 230, loss: 5.2215647883713245e-05  
step: 240, loss: 4.863911817665212e-05  
step: 250, loss: 2.9581313356175087e-05  
step: 260, loss: 1.4774312148801982e-05  
step: 270, loss: 0.006307163741439581  
step: 280, loss: 1.8376518710283563e-05  
step: 290, loss: 3.053095133509487e-05  
step: 300, loss: 1.9151182641508058e-05  
step: 310, loss: 1.9736036847461946e-05  
step: 320, loss: 8.829886792227626e-05  
step: 330, loss: 1.728499228192959e-05  
step: 340, loss: 3.762296182685532e-05  
step: 350, loss: 0.00016648885502945632  
step: 360, loss: 0.004007294308394194  
step: 370, loss: 3.412885780562647e-05  
step: 380, loss: 2.607144597277511e-05  
step: 390, loss: 0.0010142273968085647  
step: 400, loss: 3.236022166674957e-05  
step: 410, loss: 1.86483721336117e-05  
step: 420, loss: 1.595520916453097e-05  
step: 430, loss: 2.0067767763976008e-05  
step: 440, loss: 1.5545385394943878e-05  
step: 450, loss: 1.8976252249558456e-05  
step: 460, loss: 2.7074132958659902e-05  
step: 470, loss: 2.2660153263132088e-05  
step: 480, loss: 2.1777397705591284e-05  
step: 490, loss: 2.7948268325417303e-05  
step: 500, loss: 0.00010938107152469456  
step: 510, loss: 8.026082650758326e-05  
step: 520, loss: 5.505133594851941e-05  
step: 530, loss: 2.5763225494301878e-05

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning:  
An input tensor was not cuda.

warnings.warn("An input tensor was not cuda.")  
epoch 14: dev\_f1=0.9168646080760096, f1=0.916431924882629,  
best\_f1=0.9252336448598131  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning:  
An input tensor was not cuda.

warnings.warn("An input tensor was not cuda.")

step: 0, loss: 1.816773510654457e-05  
step: 10, loss: 4.8538739065406844e-05  
step: 20, loss: 4.069367787451483e-05  
step: 30, loss: 0.00023880813387222588  
step: 40, loss: 1.5683173842262477e-05  
step: 50, loss: 6.833984662080184e-05  
step: 60, loss: 1.561238786962349e-05  
step: 70, loss: 0.061328284442424774  
step: 80, loss: 6.430810026358813e-05  
step: 90, loss: 1.6495225281687453e-05  
step: 100, loss: 2.318911356269382e-05  
step: 110, loss: 0.021404040977358818  
step: 120, loss: 0.00024631372070871294  
step: 130, loss: 1.5225000424834434e-05  
step: 140, loss: 2.439471245452296e-05  
step: 150, loss: 2.287579445692245e-05  
step: 160, loss: 5.4494252253789455e-05  
step: 170, loss: 4.980088124284521e-05  
step: 180, loss: 2.8906197258038446e-05  
step: 190, loss: 6.101692997617647e-05  
step: 200, loss: 2.284285073983483e-05  
step: 210, loss: 0.0001473143056500703  
step: 220, loss: 1.3258174476504792e-05  
step: 230, loss: 0.0026483142282813787  
step: 240, loss: 2.3319791580433957e-05  
step: 250, loss: 0.0001571802276885137  
step: 260, loss: 4.034379162476398e-05  
step: 270, loss: 0.0024921686854213476  
step: 280, loss: 6.362266867654398e-05  
step: 290, loss: 0.00030641769990324974  
step: 300, loss: 0.0005247212829999626  
step: 310, loss: 0.0003292171750217676  
step: 320, loss: 2.82402088487288e-05  
step: 330, loss: 3.6471108614932746e-05  
step: 340, loss: 3.6757603083970025e-05  
step: 350, loss: 2.419845804979559e-05  
step: 360, loss: 6.412401853594929e-05  
step: 370, loss: 2.146062979591079e-05  
step: 380, loss: 7.405632641166449e-05

```

step: 390, loss: 0.00011058132076868787
step: 400, loss: 1.5914181858533993e-05
step: 410, loss: 2.4828279492794536e-05
step: 420, loss: 2.490649785613641e-05
step: 430, loss: 2.4235914679593407e-05
step: 440, loss: 2.9197699404903688e-05
step: 450, loss: 0.0003218149358872324
step: 460, loss: 1.2930333468830213e-05
step: 470, loss: 5.520370905287564e-05
step: 480, loss: 1.5508128853980452e-05
step: 490, loss: 2.1580122847808525e-05
step: 500, loss: 5.914569192100316e-05
step: 510, loss: 0.0014712824486196041
step: 520, loss: 2.6258188881911337e-05
step: 530, loss: 3.0117280402919278e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.918532634587899, f1=0.9197355996222852,
best_f1=0.9252336448598131
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-

```



```

packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
5742it [00:13, 424.59it/s]
load_f1 = 0.928772258669166
real_f1 = 0.9273066169617894
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:10, 423.02it/s]

```

#### 1.1.4 Amazon-Google - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/Amazon-Google \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm distilbert \
  --fp16 \
  --save_model \
  --da entry_swap \
  --da span_del \
  --summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Amazon-Google.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/Amazon-Google \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm distilbert \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback.

Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

warnings.warn("An input tensor was not cuda.")

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0  
/usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`.

In PyTorch 1.1.0 and later, you should call them in the opposite order:

`optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

"<https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>",  
UserWarning)

step: 0, loss: 0.8323044776916504

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

step: 10, loss: 0.06231743469834328

step: 20, loss: 0.37183335423469543

step: 30, loss: 0.37321022152900696

step: 40, loss: 0.4994754493236542

step: 50, loss: 0.2924441993236542

step: 60, loss: 0.3647793233394623

step: 70, loss: 0.2153022438287735

step: 80, loss: 0.27937808632850647

step: 90, loss: 0.4012330174446106

step: 100, loss: 0.19889244437217712

step: 110, loss: 0.2910327911376953

step: 120, loss: 0.23252302408218384

step: 130, loss: 0.25045523047447205

```

step: 140, loss: 0.23907075822353363
step: 150, loss: 0.29802611470222473
step: 160, loss: 0.16269676387310028
step: 170, loss: 0.12736515700817108
step: 180, loss: 0.19374597072601318
step: 190, loss: 0.17529296875
step: 200, loss: 0.1936153769493103
step: 210, loss: 0.4885278046131134
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.5089605734767024, f1=0.5392857142857143,
best_f1=0.5392857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0887300968170166
step: 10, loss: 0.06552992016077042
step: 20, loss: 0.2774491608142853
step: 30, loss: 0.11424334347248077
step: 40, loss: 0.11905153095722198
step: 50, loss: 0.16519927978515625
step: 60, loss: 0.027152257040143013
step: 70, loss: 0.14815542101860046
step: 80, loss: 0.24312211573123932
step: 90, loss: 0.12593431770801544
step: 100, loss: 0.10462846606969833
step: 110, loss: 0.08139154314994812
step: 120, loss: 0.2892187237739563
step: 130, loss: 0.2354261577129364
step: 140, loss: 0.23140935599803925
step: 150, loss: 0.23309046030044556
step: 160, loss: 0.11158116161823273
step: 170, loss: 0.2138616293668747
step: 180, loss: 0.2909691631793976
step: 190, loss: 0.20240727066993713
step: 200, loss: 0.18472173810005188
step: 210, loss: 0.402334988117218
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.

```

```

    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.5736925515055469, f1=0.5768025078369905,
best_f1=0.5768025078369905
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20974424481391907
step: 10, loss: 0.2699137032032013
step: 20, loss: 0.2783990204334259
step: 30, loss: 0.06497743725776672
step: 40, loss: 0.17061588168144226
step: 50, loss: 0.10686022788286209
step: 60, loss: 0.15665383636951447
step: 70, loss: 0.12675519287586212
step: 80, loss: 0.07021483033895493
step: 90, loss: 0.2389136105775833
step: 100, loss: 0.024456115439534187
step: 110, loss: 0.2024294137954712
step: 120, loss: 0.20329159498214722
step: 130, loss: 0.16295693814754486
step: 140, loss: 0.29852497577667236
step: 150, loss: 0.27868056297302246
step: 160, loss: 0.0890849232673645
step: 170, loss: 0.14949049055576324
step: 180, loss: 0.06975267082452774
step: 190, loss: 0.19823385775089264
step: 200, loss: 0.19913016259670258
step: 210, loss: 0.2361462414264679
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.5886939571150098, f1=0.6165703275529865,
best_f1=0.6165703275529865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1647413671016693
step: 10, loss: 0.05352632701396942
step: 20, loss: 0.19219842553138733
step: 30, loss: 0.051034461706876755
step: 40, loss: 0.14513008296489716
step: 50, loss: 0.11455734074115753
step: 60, loss: 0.13265585899353027
step: 70, loss: 0.20135578513145447

```

```

step: 80, loss: 0.18196862936019897
step: 90, loss: 0.012008005753159523
step: 100, loss: 0.06522537022829056
step: 110, loss: 0.17543238401412964
step: 120, loss: 0.04000179097056389
step: 130, loss: 0.3174458146095276
step: 140, loss: 0.16222061216831207
step: 150, loss: 0.10718531906604767
step: 160, loss: 0.0667974054813385
step: 170, loss: 0.048698365688323975
step: 180, loss: 0.16776788234710693
step: 190, loss: 0.08278881758451462
step: 200, loss: 0.20141199231147766
step: 210, loss: 0.04478561878204346
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.5978260869565217, f1=0.6171003717472119,
best_f1=0.6171003717472119
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19746938347816467
step: 10, loss: 0.10748089849948883
step: 20, loss: 0.04308177903294563
step: 30, loss: 0.09250741451978683
step: 40, loss: 0.08902425318956375
step: 50, loss: 0.2850453853607178
step: 60, loss: 0.06343735754489899
step: 70, loss: 0.07031228393316269
step: 80, loss: 0.05825797840952873
step: 90, loss: 0.06604088097810745
step: 100, loss: 0.047436438500881195
step: 110, loss: 0.007399728987365961
step: 120, loss: 0.022130683064460754
step: 130, loss: 0.13230088353157043
step: 140, loss: 0.026058681309223175
step: 150, loss: 0.23878920078277588
step: 160, loss: 0.017273591831326485
step: 170, loss: 0.057981155812740326
step: 180, loss: 0.16162286698818207
step: 190, loss: 0.14630495011806488
step: 200, loss: 0.10869187861680984
step: 210, loss: 0.056706808507442474

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.5754527162977867, f1=0.5979797979797981,
best_f1=0.6171003717472119
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06036679819226265
step: 10, loss: 0.03112032450735569
step: 20, loss: 0.06558290868997574
step: 30, loss: 0.15616488456726074
step: 40, loss: 0.05781059339642525
step: 50, loss: 0.12032239139080048
step: 60, loss: 0.19325359165668488
step: 70, loss: 0.03324956074357033
step: 80, loss: 0.10981934517621994
step: 90, loss: 0.11571528017520905
step: 100, loss: 0.031305525451898575
step: 110, loss: 0.08717348426580429
step: 120, loss: 0.022924436256289482
step: 130, loss: 0.05167918652296066
step: 140, loss: 0.0938195139169693
step: 150, loss: 0.04252054542303085
step: 160, loss: 0.2224481701850891
step: 170, loss: 0.03161875903606415
step: 180, loss: 0.011436264030635357
step: 190, loss: 0.058904971927404404
step: 200, loss: 0.06644642353057861
step: 210, loss: 0.13951951265335083
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.5593869731800766, f1=0.5633802816901408,
best_f1=0.6171003717472119
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05732140317559242
step: 10, loss: 0.042397335171699524
step: 20, loss: 0.03404633700847626
step: 30, loss: 0.04567863792181015
step: 40, loss: 0.043678246438503265
step: 50, loss: 0.04332534223794937
step: 60, loss: 0.21438701450824738
step: 70, loss: 0.006270186509937048
step: 80, loss: 0.10404578596353531
step: 90, loss: 0.058025091886520386

```

```

step: 100, loss: 0.029927629977464676
step: 110, loss: 0.1663009226322174
step: 120, loss: 0.26508381962776184
step: 130, loss: 0.025744890794157982
step: 140, loss: 0.010220462456345558
step: 150, loss: 0.05735066905617714
step: 160, loss: 0.08166208863258362
step: 170, loss: 0.054488398134708405
step: 180, loss: 0.004329612944275141
step: 190, loss: 0.041580893099308014
step: 200, loss: 0.11639387905597687
step: 210, loss: 0.02126152440905571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.5863453815261044, f1=0.5945945945945946,
best_f1=0.6171003717472119
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.038811083883047104
step: 10, loss: 0.030570894479751587
step: 20, loss: 0.009500083513557911
step: 30, loss: 0.03688795492053032
step: 40, loss: 0.03287205472588539
step: 50, loss: 0.09271679073572159
step: 60, loss: 0.3626234829425812
step: 70, loss: 0.006521687842905521
step: 80, loss: 0.04295768961310387
step: 90, loss: 0.007412922568619251
step: 100, loss: 0.008905204012989998
step: 110, loss: 0.016215022653341293
step: 120, loss: 0.010985399596393108
step: 130, loss: 0.015141835436224937
step: 140, loss: 0.009275652468204498
step: 150, loss: 0.05703245848417282
step: 160, loss: 0.041345927864313126
step: 170, loss: 0.03704944625496864
step: 180, loss: 0.03864813968539238
step: 190, loss: 0.005149708595126867
step: 200, loss: 0.025627830997109413
step: 210, loss: 0.16424617171287537
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.5714285714285715, f1=0.5560165975103734,
best_f1=0.6171003717472119
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.042153067886829376
step: 10, loss: 0.048134639859199524
step: 20, loss: 0.05891750007867813
step: 30, loss: 0.019810091704130173
step: 40, loss: 0.09233145415782928
step: 50, loss: 0.10458178073167801
step: 60, loss: 0.016754938289523125
step: 70, loss: 0.05064065009355545
step: 80, loss: 0.02835489995777607
step: 90, loss: 0.07759716361761093
step: 100, loss: 0.024755654856562614
step: 110, loss: 0.020970571786165237
step: 120, loss: 0.07872045785188675
step: 130, loss: 0.022150276228785515
step: 140, loss: 0.029294539242982864
step: 150, loss: 0.0703209862112999
step: 160, loss: 0.010891471058130264
step: 170, loss: 0.01639564149081707
step: 180, loss: 0.06519503891468048
step: 190, loss: 0.180546373128891
step: 200, loss: 0.01415608823299408
step: 210, loss: 0.17584048211574554
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.5689948892674618, f1=0.5540069686411151,
best_f1=0.6171003717472119
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008995131589472294
step: 10, loss: 0.0075461347587406635
step: 20, loss: 0.033240657299757004
step: 30, loss: 0.023445580154657364
step: 40, loss: 0.07507415860891342
step: 50, loss: 0.013758419081568718
step: 60, loss: 0.002080832375213504
step: 70, loss: 0.10173245519399643
step: 80, loss: 0.0007521893130615354
step: 90, loss: 0.059772174805402756
step: 100, loss: 0.0013011816190555692
step: 110, loss: 0.003930804785341024
step: 120, loss: 0.016813797876238823
step: 130, loss: 0.0018841709243133664
step: 140, loss: 0.003707454539835453
step: 150, loss: 0.06753800809383392

```



```

step: 160, loss: 0.1262696236371994
step: 170, loss: 0.010483911260962486
step: 180, loss: 0.0038809129036962986
step: 190, loss: 0.28708526492118835
step: 200, loss: 0.024523474276065826
step: 210, loss: 0.017673857510089874
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.5782312925170068, f1=0.5719424460431655,
best_f1=0.6171003717472119
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.031962744891643524
step: 10, loss: 0.013740158639848232
step: 20, loss: 0.008374519646167755
step: 30, loss: 0.007714494597166777
step: 40, loss: 0.023868095129728317
step: 50, loss: 0.010746094398200512
step: 60, loss: 0.036689501255750656
step: 70, loss: 0.02547161653637886
step: 80, loss: 0.16565540432929993
step: 90, loss: 0.2065182328224182
step: 100, loss: 0.08379105478525162
step: 110, loss: 0.006683360785245895
step: 120, loss: 0.0033270104322582483
step: 130, loss: 0.1072021946310997
step: 140, loss: 0.03059215657413006
step: 150, loss: 0.18174278736114502
step: 160, loss: 0.026903579011559486
step: 170, loss: 0.043865010142326355
step: 180, loss: 0.01570233702659607
step: 190, loss: 0.027955597266554832
step: 200, loss: 0.00402247766032815
step: 210, loss: 0.055497609078884125
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.5709624796084829, f1=0.5543859649122808,
best_f1=0.6171003717472119
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021900273859500885
step: 10, loss: 0.07187315821647644
step: 20, loss: 0.013604333624243736
step: 30, loss: 0.00565275177359581

```

```
step: 40, loss: 0.1377718597650528
step: 50, loss: 0.014056227169930935
step: 60, loss: 0.0059947180561721325
step: 70, loss: 0.004333795513957739
step: 80, loss: 0.004942889325320721
step: 90, loss: 0.0012616415042430162
step: 100, loss: 0.09160996228456497
step: 110, loss: 0.05425667017698288
step: 120, loss: 0.013668875209987164
step: 130, loss: 0.10696878284215927
step: 140, loss: 0.010125724598765373
step: 150, loss: 0.011679105460643768
step: 160, loss: 0.015558400191366673
step: 170, loss: 0.018126318231225014
step: 180, loss: 0.0075036766938865185
step: 190, loss: 0.008814643137156963
step: 200, loss: 0.08928190916776657
step: 210, loss: 0.03483331575989723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.5804066543438077, f1=0.5802707930367504,
best_f1=0.6171003717472119
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014809982385486364
step: 10, loss: 0.00604881951585412
step: 20, loss: 0.0016474996227771044
step: 30, loss: 0.018001705408096313
step: 40, loss: 0.001022079843096435
step: 50, loss: 0.005667218938469887
step: 60, loss: 0.009296159259974957
step: 70, loss: 0.07520318031311035
step: 80, loss: 0.01435500755906105
step: 90, loss: 0.06657830625772476
step: 100, loss: 0.15411819517612457
step: 110, loss: 0.013356370851397514
step: 120, loss: 0.08482448756694794
step: 130, loss: 0.0012841739226132631
step: 140, loss: 0.015501783229410648
step: 150, loss: 0.009214293211698532
step: 160, loss: 0.14783349633216858
step: 170, loss: 0.007546961773186922
step: 180, loss: 0.07467233389616013
step: 190, loss: 0.003058433998376131
step: 200, loss: 0.004419444594532251
step: 210, loss: 0.010195179842412472
```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.5693430656934306, f1=0.5725047080979284,
best_f1=0.6171003717472119
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013738983310759068
step: 10, loss: 0.004886308219283819
step: 20, loss: 0.013966110534965992
step: 30, loss: 0.025970427319407463
step: 40, loss: 0.016380269080400467
step: 50, loss: 0.013899353332817554
step: 60, loss: 0.06924471259117126
step: 70, loss: 0.014169973321259022
step: 80, loss: 0.13156849145889282
step: 90, loss: 0.10768191516399384
step: 100, loss: 0.006763352081179619
step: 110, loss: 0.023518815636634827
step: 120, loss: 0.008974636904895306
step: 130, loss: 0.008250080980360508
step: 140, loss: 0.0014382473891600966
step: 150, loss: 0.02696213871240616
step: 160, loss: 0.14900057017803192
step: 170, loss: 0.0048187547363340855
step: 180, loss: 0.0036170792300254107
step: 190, loss: 0.011885935440659523
step: 200, loss: 0.0025066290982067585
step: 210, loss: 0.00666449312120676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.5725338491295938, f1=0.5708502024291497,
best_f1=0.6171003717472119
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0546974316239357
step: 10, loss: 0.016785820946097374
step: 20, loss: 0.0066711679100990295
step: 30, loss: 0.0004364675551187247
step: 40, loss: 0.0007212362252175808
step: 50, loss: 0.0019766746554523706
step: 60, loss: 0.04776414483785629
step: 70, loss: 0.003637452144175768
step: 80, loss: 0.0030168206430971622
step: 90, loss: 0.004416537471115589

```

```

step: 100, loss: 0.0037150231655687094
step: 110, loss: 0.006721393670886755
step: 120, loss: 0.07001131772994995
step: 130, loss: 0.04369325190782547
step: 140, loss: 0.009224908426404
step: 150, loss: 0.002343036700040102
step: 160, loss: 0.02009207382798195
step: 170, loss: 0.005973674822598696
step: 180, loss: 0.04355306923389435
step: 190, loss: 0.0017487803706899285
step: 200, loss: 0.01719105802476406
step: 210, loss: 0.03702687472105026
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.5764023210831721, f1=0.5680000000000001,
best_f1=0.6171003717472119
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2293it [00:06, 343.40it/s]

```

```
load_f1 = 0.598360655737705
real_f1 = 0.5967078189300411
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:17, 254.66it/s]
```

### 1.1.5 Walmart-Amazon - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/Walmart-Amazon \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm distilbert \
  --fp16 \
  --save_model \
  --da entry_swap \
  --da span_del \
  --summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Walmart-Amazon.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/Walmart-Amazon \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm distilbert \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/
```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```
enabled          : True
opt_level        : 02
```

```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled              : True
opt_level            : 02
cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.8627281188964844
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.1652759313583374
step: 20, loss: 0.15136198699474335
step: 30, loss: 0.5041339993476868
step: 40, loss: 0.25074997544288635
step: 50, loss: 0.30944180488586426
step: 60, loss: 0.36331480741500854
step: 70, loss: 0.17748676240444183
step: 80, loss: 0.5313534736633301
step: 90, loss: 0.24622014164924622
step: 100, loss: 0.21772751212120056
step: 110, loss: 0.22917908430099487
step: 120, loss: 0.40507271885871887
step: 130, loss: 0.35548314452171326
step: 140, loss: 0.2858782112598419
step: 150, loss: 0.2812540829181671
step: 160, loss: 0.2200714349746704

```

```

step: 170, loss: 0.36758580803871155
step: 180, loss: 0.27844005823135376
step: 190, loss: 0.14471030235290527
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.5801526717557252, f1=0.6203473945409429,
best_f1=0.6203473945409429
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23293974995613098
step: 10, loss: 0.04068192094564438
step: 20, loss: 0.03857199847698212
step: 30, loss: 0.1298290342092514
step: 40, loss: 0.4123183786869049
step: 50, loss: 0.3329295217990875
step: 60, loss: 0.11436765640974045
step: 70, loss: 0.18747979402542114
step: 80, loss: 0.215900719165802
step: 90, loss: 0.13192759454250336
step: 100, loss: 0.29275715351104736
step: 110, loss: 0.14025650918483734
step: 120, loss: 0.3371172547340393
step: 130, loss: 0.17414435744285583
step: 140, loss: 0.24099504947662354
step: 150, loss: 0.04099597781896591
step: 160, loss: 0.04156967252492905
step: 170, loss: 0.14655248820781708
step: 180, loss: 0.16925343871116638
step: 190, loss: 0.18656407296657562
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.7393617021276595, f1=0.7539267015706806,
best_f1=0.7539267015706806
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1507781445980072

```

```

step: 10, loss: 0.23624271154403687
step: 20, loss: 0.07458420842885971
step: 30, loss: 0.04757536202669144
step: 40, loss: 0.03667478635907173
step: 50, loss: 0.06200757622718811
step: 60, loss: 0.01033977884799242
step: 70, loss: 0.0733000636100769
step: 80, loss: 0.20300962030887604
step: 90, loss: 0.042663637548685074
step: 100, loss: 0.08028480410575867
step: 110, loss: 0.29109877347946167
step: 120, loss: 0.031148690730333328
step: 130, loss: 0.08639239519834518
step: 140, loss: 0.12072663009166718
step: 150, loss: 0.19363224506378174
step: 160, loss: 0.0981975570321083
step: 170, loss: 0.04782075434923172
step: 180, loss: 0.006381503771990538
step: 190, loss: 0.06496629118919373
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.736842105263158, f1=0.7714285714285715,
best_f1=0.7539267015706806
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014619159512221813
step: 10, loss: 0.08823398500680923
step: 20, loss: 0.011225046589970589
step: 30, loss: 0.04357690364122391
step: 40, loss: 0.004361575935035944
step: 50, loss: 0.06483972817659378
step: 60, loss: 0.23260167241096497
step: 70, loss: 0.015371773391962051
step: 80, loss: 0.024504324421286583
step: 90, loss: 0.05252484232187271
step: 100, loss: 0.04739505052566528
step: 110, loss: 0.09723280370235443
step: 120, loss: 0.13007166981697083
step: 130, loss: 0.22942090034484863
step: 140, loss: 0.11450473219156265
step: 150, loss: 0.010812715627253056
step: 160, loss: 0.011057642288506031
step: 170, loss: 0.07960354536771774
step: 180, loss: 0.058647021651268005
step: 190, loss: 0.0466674342751503
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```



```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.7534626038781163, f1=0.7790055248618786,
best_f1=0.7790055248618786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16047996282577515
step: 10, loss: 0.010277354158461094
step: 20, loss: 0.07284490019083023
step: 30, loss: 0.04390813410282135
step: 40, loss: 0.11084466427564621
step: 50, loss: 0.07757872343063354
step: 60, loss: 0.019390659406781197
step: 70, loss: 0.01699581928551197
step: 80, loss: 0.032181255519390106
step: 90, loss: 0.003373864572495222
step: 100, loss: 0.116301529109478
step: 110, loss: 0.06509976834058762
step: 120, loss: 0.0022546472027897835
step: 130, loss: 0.0292314812541008
step: 140, loss: 0.0050976392813026905
step: 150, loss: 0.04624101147055626
step: 160, loss: 0.006070702336728573
step: 170, loss: 0.09843385219573975
step: 180, loss: 0.12923066318035126
step: 190, loss: 0.30080074071884155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.726775956284153, f1=0.7450980392156863,
best_f1=0.7790055248618786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006644912995398045
step: 10, loss: 0.0032887342385947704
step: 20, loss: 0.009697400964796543
step: 30, loss: 0.006220496259629726
step: 40, loss: 0.14612004160881042
step: 50, loss: 0.016898753121495247
step: 60, loss: 0.008747248910367489
step: 70, loss: 0.06729578226804733
step: 80, loss: 0.022672563791275024

```

```

step: 90, loss: 0.0015131463296711445
step: 100, loss: 0.001669836463406682
step: 110, loss: 0.029279178008437157
step: 120, loss: 0.013590537011623383
step: 130, loss: 0.013908403925597668
step: 140, loss: 0.00505764689296484
step: 150, loss: 0.028377292677760124
step: 160, loss: 0.021396417170763016
step: 170, loss: 0.06531824171543121
step: 180, loss: 0.054763536900281906
step: 190, loss: 0.08313145488500595
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.7272727272727272, f1=0.7500000000000001,
best_f1=0.7790055248618786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021498920395970345
step: 10, loss: 0.03631068393588066
step: 20, loss: 0.0015070575755089521
step: 30, loss: 0.01147669367492199
step: 40, loss: 0.04715743288397789
step: 50, loss: 0.3164388835430145
step: 60, loss: 0.003354431129992008
step: 70, loss: 0.002382120583206415
step: 80, loss: 0.011814429424703121
step: 90, loss: 0.10272406786680222
step: 100, loss: 0.007812370080500841
step: 110, loss: 0.006930799223482609
step: 120, loss: 0.005597988609224558
step: 130, loss: 0.003590402426198125
step: 140, loss: 0.003837928641587496
step: 150, loss: 0.01959672011435032
step: 160, loss: 0.0033139605075120926
step: 170, loss: 0.008710966445505619
step: 180, loss: 0.0012177402386441827
step: 190, loss: 0.005364157259464264
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.7277227272727278, f1=0.7461139896373058,
best_f1=0.7790055248618786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009227515198290348

```

```

step: 10, loss: 0.005335086956620216
step: 20, loss: 0.004775172099471092
step: 30, loss: 0.003069092985242605
step: 40, loss: 0.019133862107992172
step: 50, loss: 0.0010222936980426311
step: 60, loss: 0.002053720410913229
step: 70, loss: 0.01596858911216259
step: 80, loss: 0.017985016107559204
step: 90, loss: 0.0016965130344033241
step: 100, loss: 0.07262596487998962
step: 110, loss: 0.000832185207400471
step: 120, loss: 0.020484691485762596
step: 130, loss: 0.0034657721407711506
step: 140, loss: 0.0030142152681946754
step: 150, loss: 0.0013846774818375707
step: 160, loss: 0.002194564091041684
step: 170, loss: 0.0028538487385958433
step: 180, loss: 0.0006732389447279274
step: 190, loss: 0.009698761627078056
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.7191011235955057, f1=0.7401129943502824,
best_f1=0.7790055248618786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0032395406160503626
step: 10, loss: 0.001913963002152741
step: 20, loss: 0.11468597501516342
step: 30, loss: 0.007088850252330303
step: 40, loss: 0.031085118651390076
step: 50, loss: 0.0010899092303588986
step: 60, loss: 0.0043619005009531975
step: 70, loss: 0.0022848567459732294
step: 80, loss: 0.08441010862588882
step: 90, loss: 0.030575891956686974
step: 100, loss: 0.001520251971669495
step: 110, loss: 0.10872280597686768
step: 120, loss: 0.009622105397284031
step: 130, loss: 0.011253880336880684
step: 140, loss: 0.0065771047957241535
step: 150, loss: 0.01151640247553587
step: 160, loss: 0.005138847045600414
step: 170, loss: 0.029499145224690437
step: 180, loss: 0.018921716138720512
step: 190, loss: 0.0020158356055617332
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.7122507122507123, f1=0.7256637168141593,
best_f1=0.7790055248618786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000630105787422508
step: 10, loss: 0.002843384863808751
step: 20, loss: 0.003179053543135524
step: 30, loss: 0.19030198454856873
step: 40, loss: 0.017545808106660843
step: 50, loss: 0.005718502681702375
step: 60, loss: 0.006579682230949402
step: 70, loss: 0.012446876615285873
step: 80, loss: 0.004584795795381069
step: 90, loss: 0.03425312414765358
step: 100, loss: 0.017151962965726852
step: 110, loss: 0.04329383373260498
step: 120, loss: 0.03229783475399017
step: 130, loss: 0.0035278687719255686
step: 140, loss: 0.001607235404662788
step: 150, loss: 0.0013914534356445074
step: 160, loss: 0.19858980178833008
step: 170, loss: 0.004590020515024662
step: 180, loss: 0.0011165005853399634
step: 190, loss: 0.014343341812491417
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.7163323782234957, f1=0.7383720930232558,
best_f1=0.7790055248618786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008183865807950497
step: 10, loss: 0.002815751824527979
step: 20, loss: 0.026774220168590546
step: 30, loss: 0.002943689003586769
step: 40, loss: 0.001241572666913271
step: 50, loss: 0.0014422291424125433
step: 60, loss: 0.0005920652183704078
step: 70, loss: 0.0006571168778464198
step: 80, loss: 0.001505292602814734
step: 90, loss: 0.005751711782068014
step: 100, loss: 0.0012686968548223376
step: 110, loss: 0.0006011563236825168
step: 120, loss: 0.016402827575802803

```

```

step: 130, loss: 0.009751233272254467
step: 140, loss: 0.0010627154260873795
step: 150, loss: 0.0011193887330591679
step: 160, loss: 0.001806829241104424
step: 170, loss: 0.02734927274286747
step: 180, loss: 0.09161894023418427
step: 190, loss: 0.003471602685749531
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.7371273712737126, f1=0.7554347826086956,
best_f1=0.7790055248618786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009038177318871021
step: 10, loss: 0.03885161131620407
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.000800850335508585
step: 30, loss: 0.0013061205390840769
step: 40, loss: 0.002514971885830164
step: 50, loss: 0.002795384731143713
step: 60, loss: 0.0006583755603060126
step: 70, loss: 0.0006621509091928601
step: 80, loss: 0.0006359820836223662
step: 90, loss: 0.002459599170833826
step: 100, loss: 0.0022906239610165358
step: 110, loss: 0.0006229804130271077
step: 120, loss: 0.001993571873754263
step: 130, loss: 0.0002844429691322148
step: 140, loss: 0.0029665075708180666
step: 150, loss: 0.0013573355972766876
step: 160, loss: 0.031109875068068504
step: 170, loss: 0.03458084538578987
step: 180, loss: 0.008512596599757671
step: 190, loss: 0.03891954571008682
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.7322404371584699, f1=0.7567567567567567,
best_f1=0.7790055248618786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09938324242830276
step: 10, loss: 0.029128266498446465
step: 20, loss: 0.003894915571436286
step: 30, loss: 0.005139201879501343

```

```

step: 40, loss: 0.0005507831228896976
step: 50, loss: 0.007333869580179453
step: 60, loss: 0.0008786170510575175
step: 70, loss: 0.0018157983431592584
step: 80, loss: 0.0004528590361587703
step: 90, loss: 0.0008611868834123015
step: 100, loss: 0.03015756420791149
step: 110, loss: 0.007257347460836172
step: 120, loss: 0.015561102889478207
step: 130, loss: 0.001160676241852343
step: 140, loss: 0.008371347561478615
step: 150, loss: 0.0005784022505395114
step: 160, loss: 0.0026473393663764
step: 170, loss: 0.0008932322962209582
step: 180, loss: 0.009465189650654793
step: 190, loss: 0.0736672431230545
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.7388888888888889, f1=0.7478753541076487,
best_f1=0.7790055248618786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.027212142944335938
step: 10, loss: 0.0005243188352324069
step: 20, loss: 0.0014418866485357285
step: 30, loss: 0.0026641390286386013
step: 40, loss: 0.004141032695770264
step: 50, loss: 0.021691102534532547
step: 60, loss: 0.0011141895083710551
step: 70, loss: 0.001964314142242074
step: 80, loss: 0.0005199351580813527
step: 90, loss: 0.009100314229726791
step: 100, loss: 0.0009138141176663339
step: 110, loss: 0.07035955786705017
step: 120, loss: 0.09024832397699356
step: 130, loss: 0.0007841189508326352
step: 140, loss: 0.0010029837721958756
step: 150, loss: 0.0017079152166843414
step: 160, loss: 0.0006878937128931284
step: 170, loss: 0.003746490925550461
step: 180, loss: 0.004319400526583195
step: 190, loss: 0.00017621871666051447
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.732394366197183, f1=0.7619047619047619,

```

```

best_f1=0.7790055248618786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014180241851136088
step: 10, loss: 0.0006986148655414581
step: 20, loss: 0.004189381841570139
step: 30, loss: 0.0015023304149508476
step: 40, loss: 0.005530897062271833
step: 50, loss: 0.021563084796071053
step: 60, loss: 0.003216851269826293
step: 70, loss: 0.0009229768184013665
step: 80, loss: 0.00044475324102677405
step: 90, loss: 0.007384082302451134
step: 100, loss: 0.00039354764157906175
step: 110, loss: 0.0013323985040187836
step: 120, loss: 0.0006939179729670286
step: 130, loss: 0.000703567115124315
step: 140, loss: 0.004755575675517321
step: 150, loss: 0.0007637051166966558
step: 160, loss: 0.009683586657047272
step: 170, loss: 0.0016443687491118908
step: 180, loss: 0.003611014224588871
step: 190, loss: 0.04843718931078911
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.7344632768361581, f1=0.7627118644067796,
best_f1=0.7790055248618786
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True

```

```

master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:09, 222.00it/s]
load_f1 = 0.5829145728643216
real_f1 = 0.5817307692307693
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:17, 244.97it/s]

```

## 1.2 DA TEXTUAL

### 1.2.1 Abt-Buy - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
    --task Textual/Abt-Buy \
    --batch_size 32 \
    --max_len 256 \
    --lr 3e-5 \
    --n_epochs 15 \
    --finetuning \
    --lm distilbert \
    --fp16 \
    --save_model \
    --da entry_swap \
    --da span_shuffle \
    --summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Abt-Buy.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \

```



```

--task Textual/Abt-Buy \
--input_path input/input_small.jsonl \
--output_path output/output_small.jsonl \
--lm distilbert \
--use_gpu \
--fp16 \
--checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

warnings.warn("An input tensor was not cuda.")

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0  
/usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`.

In PyTorch 1.1.0 and later, you should call them in the opposite order: `optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

"https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",  
UserWarning)

step: 0, loss: 0.8541277050971985

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0  
step: 10, loss: 0.22715064883232117

```

step: 20, loss: 0.14705081284046173
step: 30, loss: 0.2407722920179367
step: 40, loss: 0.31026098132133484
step: 50, loss: 0.38548168540000916
step: 60, loss: 0.4559246301651001
step: 70, loss: 0.3066379129886627
step: 80, loss: 0.24344772100448608
step: 90, loss: 0.4001930058002472
step: 100, loss: 0.22453896701335907
step: 110, loss: 0.1813557893037796
step: 120, loss: 0.5387949347496033
step: 130, loss: 0.4022427797317505
step: 140, loss: 0.4463738799095154
step: 150, loss: 0.06160173937678337
step: 160, loss: 0.2215418517589569
step: 170, loss: 0.14646612107753754
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.6077097505668935, f1=0.6463700234192037,
best_f1=0.6463700234192037
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2907020151615143
step: 10, loss: 0.1748431771993637
step: 20, loss: 0.2514936625957489
step: 30, loss: 0.3739021420478821
step: 40, loss: 0.16291648149490356
step: 50, loss: 0.2066945880651474
step: 60, loss: 0.09434592723846436
step: 70, loss: 0.18574398756027222
step: 80, loss: 0.08479534089565277
step: 90, loss: 0.15805326402187347
step: 100, loss: 0.13677208125591278
step: 110, loss: 0.16769662499427795
step: 120, loss: 0.056273266673088074
step: 130, loss: 0.057331718504428864
step: 140, loss: 0.13240982592105865
step: 150, loss: 0.21663829684257507
step: 160, loss: 0.13431018590927124
step: 170, loss: 0.13729329407215118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.7241379310344829, f1=0.7614457831325301,
best_f1=0.7614457831325301
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0225977823138237
step: 10, loss: 0.05924651399254799
step: 20, loss: 0.005864674225449562
step: 30, loss: 0.17162330448627472
step: 40, loss: 0.015694204717874527
step: 50, loss: 0.12514032423496246
step: 60, loss: 0.04924117773771286
step: 70, loss: 0.04292214289307594
step: 80, loss: 0.1884026676416397
step: 90, loss: 0.07186748832464218
step: 100, loss: 0.030322598293423653
step: 110, loss: 0.08055394887924194
step: 120, loss: 0.011099238879978657
step: 130, loss: 0.11328838020563126
step: 140, loss: 0.017207089811563492
step: 150, loss: 0.08104829490184784
step: 160, loss: 0.038649629801511765
step: 170, loss: 0.10621128976345062
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.7680412371134021, f1=0.801007556675063,
best_f1=0.801007556675063
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04360322654247284
step: 10, loss: 0.06008409708738327
step: 20, loss: 0.04934939369559288
step: 30, loss: 0.02926081232726574
step: 40, loss: 0.012749402783811092
step: 50, loss: 0.022627132013440132
step: 60, loss: 0.1594427227973938
step: 70, loss: 0.043149884790182114

```

```

step: 80, loss: 0.03249897435307503
step: 90, loss: 0.040844693779945374
step: 100, loss: 0.028235165402293205
step: 110, loss: 0.023981748148798943
step: 120, loss: 0.11587323248386383
step: 130, loss: 0.048143379390239716
step: 140, loss: 0.03882172331213951
step: 150, loss: 0.0026914242189377546
step: 160, loss: 0.05286844074726105
step: 170, loss: 0.018732067197561264
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.7584415584415585, f1=0.7892156862745098,
best_f1=0.801007556675063
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09265559911727905
step: 10, loss: 0.046461332589387894
step: 20, loss: 0.07225542515516281
step: 30, loss: 0.07964017242193222
step: 40, loss: 0.04348771274089813
step: 50, loss: 0.012186652049422264
step: 60, loss: 0.0009706859709694982
step: 70, loss: 0.08544300496578217
step: 80, loss: 0.00991576723754406
step: 90, loss: 0.012658586725592613
step: 100, loss: 0.02099435217678547
step: 110, loss: 0.03575170785188675
step: 120, loss: 0.05722880735993385
step: 130, loss: 0.005083986092358828
step: 140, loss: 0.18743911385536194
step: 150, loss: 0.021037321537733078
step: 160, loss: 0.0371435284614563
step: 170, loss: 0.07947360724210739
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.7760416666666666, f1=0.8157248157248157,
best_f1=0.8157248157248157
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.05314837396144867
step: 10, loss: 0.10446396470069885
step: 20, loss: 0.003605913370847702
step: 30, loss: 0.03705013915896416
step: 40, loss: 0.01431665476411581
step: 50, loss: 0.06935682147741318
step: 60, loss: 0.03416023775935173
step: 70, loss: 0.12370844930410385
step: 80, loss: 0.10686373710632324
step: 90, loss: 0.0330655500292778
step: 100, loss: 0.003867761231958866
step: 110, loss: 0.024473819881677628
step: 120, loss: 0.020268386229872704
step: 130, loss: 0.21712948381900787
step: 140, loss: 0.022395765408873558
step: 150, loss: 0.21306112408638
step: 160, loss: 0.007973086088895798
step: 170, loss: 0.003440950997173786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.7371273712737128, f1=0.7989556135770236,
best_f1=0.8157248157248157
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015203750692307949
step: 10, loss: 0.0018496682168915868
step: 20, loss: 0.0810859277844429
step: 30, loss: 0.003876086790114641
step: 40, loss: 0.05708030238747597
step: 50, loss: 0.004744953475892544
step: 60, loss: 0.10279668122529984
step: 70, loss: 0.010823985561728477
step: 80, loss: 0.12346868962049484
step: 90, loss: 0.004797067493200302
step: 100, loss: 0.003240977181121707
step: 110, loss: 0.05943368002772331
step: 120, loss: 0.08396736532449722
step: 130, loss: 0.03375754505395889
step: 140, loss: 0.041543230414390564
step: 150, loss: 0.07377977669239044
step: 160, loss: 0.025040078908205032
step: 170, loss: 0.0320521742105484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.753315649867374, f1=0.810126582278481,

```

```

best_f1=0.8157248157248157
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.026261424645781517
step: 10, loss: 0.01995941251516342
step: 20, loss: 0.007064599543809891
step: 30, loss: 0.17264658212661743
step: 40, loss: 0.038437437266111374
step: 50, loss: 0.0029524601995944977
step: 60, loss: 0.006089533679187298
step: 70, loss: 0.020331833511590958
step: 80, loss: 0.000639700738247484
step: 90, loss: 0.04600151628255844
step: 100, loss: 0.015897255390882492
step: 110, loss: 0.05008034408092499
step: 120, loss: 0.00042480925912968814
step: 130, loss: 0.006056041456758976
step: 140, loss: 0.0037212390452623367
step: 150, loss: 0.06136947125196457
step: 160, loss: 0.04136057570576668
step: 170, loss: 0.004231639206409454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.7621483375959078, f1=0.8157248157248157,
best_f1=0.8157248157248157
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005770877934992313
step: 10, loss: 0.01419783290475607
step: 20, loss: 0.007487551309168339
step: 30, loss: 0.042396917939186096
step: 40, loss: 0.01259844470769167
step: 50, loss: 0.0024144307244569063
step: 60, loss: 0.0011112692300230265
step: 70, loss: 0.006662001367658377
step: 80, loss: 0.019946619868278503
step: 90, loss: 0.02590649574995041
step: 100, loss: 0.02333253063261509
step: 110, loss: 0.0004892179858870804
step: 120, loss: 0.02036808617413044
step: 130, loss: 0.0025335748214274645
step: 140, loss: 0.0010865419171750546
step: 150, loss: 0.014655577950179577
step: 160, loss: 0.12161378562450409
step: 170, loss: 0.047840990126132965

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.7575757575757577, f1=0.8129675810473815,
best_f1=0.8157248157248157
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003366682429611683
step: 10, loss: 0.13829170167446136
step: 20, loss: 0.0010685492306947708
step: 30, loss: 0.000371885544154793
step: 40, loss: 0.01850670948624611
step: 50, loss: 0.013612495735287666
step: 60, loss: 0.0014208662323653698
step: 70, loss: 0.0034967532847076654
step: 80, loss: 0.028069384396076202
step: 90, loss: 0.05002966523170471
step: 100, loss: 0.03179479017853737
step: 110, loss: 0.014354427345097065
step: 120, loss: 0.03864970803260803
step: 130, loss: 0.025982249528169632
step: 140, loss: 0.12013236433267593
step: 150, loss: 0.000959066383074969
step: 160, loss: 0.0013654979411512613
step: 170, loss: 0.16488896310329437
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.7650273224043717, f1=0.796875, best_f1=0.8157248157248157
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010877777822315693
step: 10, loss: 0.012724817730486393
step: 20, loss: 0.008135084994137287
step: 30, loss: 0.0007122073438949883
step: 40, loss: 0.006317740771919489
step: 50, loss: 0.09544225037097931
step: 60, loss: 0.06023700162768364
step: 70, loss: 0.015417062677443027
step: 80, loss: 0.0008107278263196349
step: 90, loss: 0.0032282560132443905
step: 100, loss: 0.0016695240046828985
step: 110, loss: 0.004325141664594412
step: 120, loss: 0.0028360956348478794
step: 130, loss: 0.0037256116047501564
step: 140, loss: 0.02414335124194622

```

```

step: 150, loss: 0.007241685409098864
step: 160, loss: 0.0067034075036644936
step: 170, loss: 0.05919548124074936
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.764857881136951, f1=0.8188585607940446,
best_f1=0.8157248157248157
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0042893169447779655
step: 10, loss: 0.004230415914207697
step: 20, loss: 0.018749359995126724
step: 30, loss: 0.0007426405209116638
step: 40, loss: 0.0056257243268191814
step: 50, loss: 0.000646064174361527
step: 60, loss: 0.00019627039728220552
step: 70, loss: 0.0223994143307209
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 80, loss: 0.03447829186916351
step: 90, loss: 0.0001337874127784744
step: 100, loss: 0.0004396573349367827
step: 110, loss: 0.0001573473564349115
step: 120, loss: 0.005807839799672365
step: 130, loss: 0.004721346311271191
step: 140, loss: 0.00015875615645200014
step: 150, loss: 0.022734636440873146
step: 160, loss: 0.0058520482853055
step: 170, loss: 0.0038992001209408045
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.7573333333333334, f1=0.8110831234256927,
best_f1=0.8157248157248157
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013636834919452667
step: 10, loss: 0.07018744200468063
step: 20, loss: 0.005043455399572849
step: 30, loss: 0.0015635645249858499
step: 40, loss: 0.0002613742253743112
step: 50, loss: 0.03644673153758049
step: 60, loss: 0.00016989705909509212
step: 70, loss: 0.02735620178282261
step: 80, loss: 0.0035437913611531258
step: 90, loss: 0.00028965892852284014

```



```

step: 100, loss: 0.00026325465296395123
step: 110, loss: 0.00029269562219269574
step: 120, loss: 0.011962857097387314
step: 130, loss: 0.008520563133060932
step: 140, loss: 0.0028698144014924765
step: 150, loss: 0.0006664729444310069
step: 160, loss: 0.000676597177516669
step: 170, loss: 0.0008578919223509729
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.7575757575757577, f1=0.8157248157248157,
best_f1=0.8157248157248157
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005487609887495637
step: 10, loss: 0.0001523060054751113
step: 20, loss: 0.04422882944345474
step: 30, loss: 0.00021929515060037374
step: 40, loss: 0.0037081849295645952
step: 50, loss: 0.10312841087579727
step: 60, loss: 0.019793067127466202
step: 70, loss: 0.0011032669572159648
step: 80, loss: 0.002783892210572958
step: 90, loss: 0.0016273438232019544
step: 100, loss: 0.0022194699849933386
step: 110, loss: 0.00023344146029558033
step: 120, loss: 0.0022686547599732876
step: 130, loss: 0.0040010423399508
step: 140, loss: 0.00029208860360085964
step: 150, loss: 0.0002310102863702923
step: 160, loss: 0.08032654225826263
step: 170, loss: 0.002051891293376684
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.7577937649880095, f1=0.7990654205607476,
best_f1=0.8157248157248157
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006366157904267311
step: 10, loss: 0.018385225906968117
step: 20, loss: 0.024253973737359047
step: 30, loss: 0.012157169170677662
step: 40, loss: 0.0002085847081616521
step: 50, loss: 0.0010609521996229887

```

```

step: 60, loss: 8.989783236756921e-05
step: 70, loss: 0.047347087413072586
step: 80, loss: 0.000739501672796905
step: 90, loss: 0.0004751371161546558
step: 100, loss: 0.0008721163030713797
step: 110, loss: 0.0004618339007720351
step: 120, loss: 0.004805273842066526
step: 130, loss: 0.001790569513104856
step: 140, loss: 0.00033236181479878724
step: 150, loss: 0.05723095312714577
step: 160, loss: 0.00010119124635821208
step: 170, loss: 0.00044462084770202637
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.7614457831325301, f1=0.8018648018648018,
best_f1=0.8157248157248157
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
1916it [00:05, 326.60it/s]

```

```
load_f1 = 0.6552462526766595
real_f1 = 0.6191446028513238
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:17, 252.45it/s]
```

## 1.3 DA DIRTY

### 1.3.1 DBLP-ACM - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Dirty/DBLP-ACM \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm distilbert \
  --fp16 \
  --save_model \
  --da entry_swap \
  --da span_shuffle \
  --summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-ACM.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Dirty/DBLP-ACM \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm distilbert \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/
```

```
Downloading: 100% 442/442 [00:00<00:00, 261kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 21.9MB/s]
Downloading: 100% 268M/268M [00:06<00:00, 44.3MB/s]
Selected optimization level O2: FP16 training with FP32 batchnorm and FP32
```

master weights.

Defaults for this optimization level are:

```
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback.

Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

warnings.warn("An input tensor was not cuda.")

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0

/usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`.

In PyTorch 1.1.0 and later, you should call them in the opposite order:

`optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

"<https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>", UserWarning)

step: 0, loss: 0.8204652667045593

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

step: 10, loss: 0.47375941276550293

step: 20, loss: 0.6065149903297424

step: 30, loss: 0.4752466678619385

step: 40, loss: 0.2930925786495209

step: 50, loss: 0.09408707916736603

step: 60, loss: 0.1325923055410385

step: 70, loss: 0.1782095730304718

step: 80, loss: 0.2418617606163025

step: 90, loss: 0.0427437350153923

step: 100, loss: 0.17252196371555328

step: 110, loss: 0.0735800638794899

```

step: 120, loss: 0.043274711817502975
step: 130, loss: 0.02709943801164627
step: 140, loss: 0.0926334485411644
step: 150, loss: 0.1316460818052292
step: 160, loss: 0.09597062319517136
step: 170, loss: 0.01846853457391262
step: 180, loss: 0.015521552413702011
step: 190, loss: 0.028303995728492737
step: 200, loss: 0.008658302947878838
step: 210, loss: 0.024089841172099113
step: 220, loss: 0.029795318841934204
step: 230, loss: 0.014609208330512047
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9547960308710033, f1=0.9570011025358324,
best_f1=0.9570011025358324
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007806639187037945
step: 10, loss: 0.0012541244504973292
step: 20, loss: 0.009240549989044666
step: 30, loss: 0.00702601158991456
step: 40, loss: 0.014595786109566689
step: 50, loss: 0.02151414379477501
step: 60, loss: 0.012716379947960377
step: 70, loss: 0.006986594758927822
step: 80, loss: 0.0026863054372370243
step: 90, loss: 0.004337087273597717
step: 100, loss: 0.10209443420171738
step: 110, loss: 0.06679652631282806
step: 120, loss: 0.01411974523216486
step: 130, loss: 0.010435529984533787
step: 140, loss: 0.31119439005851746
step: 150, loss: 0.015588156878948212
step: 160, loss: 0.031073393300175667
step: 170, loss: 0.1087133064866066
step: 180, loss: 0.017212150618433952
step: 190, loss: 0.1482229232788086
step: 200, loss: 0.00265285256318748
step: 210, loss: 0.023992715403437614
step: 220, loss: 0.0017857077764347196
step: 230, loss: 0.02356838621199131

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9653631284916202, f1=0.9698324022346367,
best_f1=0.9698324022346367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.052217330783605576
step: 10, loss: 0.06234757602214813
step: 20, loss: 0.026150060817599297
step: 30, loss: 0.006882623303681612
step: 40, loss: 0.04113847017288208
step: 50, loss: 0.009098397567868233
step: 60, loss: 0.0820266604423523
step: 70, loss: 0.004214929882436991
step: 80, loss: 0.0018816182855516672
step: 90, loss: 0.07648584246635437
step: 100, loss: 0.011179428547620773
step: 110, loss: 0.06914010643959045
step: 120, loss: 0.045212648808956146
step: 130, loss: 0.008085330948233604
step: 140, loss: 0.005562986247241497
step: 150, loss: 0.005339238326996565
step: 160, loss: 0.027883904054760933
step: 170, loss: 0.0024547993671149015
step: 180, loss: 0.0070290518924593925
step: 190, loss: 0.0022885871585458517
step: 200, loss: 0.001932482235133648
step: 210, loss: 0.024619368836283684
step: 220, loss: 0.12960176169872284
step: 230, loss: 0.058443810790777206
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.9580022701475595, f1=0.9513023782559457,
best_f1=0.9698324022346367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011065470054745674
step: 10, loss: 0.004884004592895508
step: 20, loss: 0.0009555012220516801
step: 30, loss: 0.0009552567498758435

```

```

step: 40, loss: 0.024187950417399406
step: 50, loss: 0.0037323134019970894
step: 60, loss: 0.0020189115311950445
step: 70, loss: 0.0946761965751648
step: 80, loss: 0.11278078705072403
step: 90, loss: 0.0041307141073048115
step: 100, loss: 0.0024829492904245853
step: 110, loss: 0.01494919415563345
step: 120, loss: 0.004085318651050329
step: 130, loss: 0.05453890562057495
step: 140, loss: 0.0012008289340883493
step: 150, loss: 0.0033491351641714573
step: 160, loss: 0.022569535300135612
step: 170, loss: 0.0014936429215595126
step: 180, loss: 0.11558952927589417
step: 190, loss: 0.017161086201667786
step: 200, loss: 0.0029047084972262383
step: 210, loss: 0.0026972449850291014
step: 220, loss: 0.0007590953609906137
step: 230, loss: 0.0005306145176291466
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.9620535714285715, f1=0.9546961325966851,
best_f1=0.9698324022346367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006381841376423836
step: 10, loss: 0.0006143294158391654
step: 20, loss: 0.04480268433690071
step: 30, loss: 0.00023730624525342137
step: 40, loss: 0.002821363043040037
step: 50, loss: 0.0008051145705394447
step: 60, loss: 0.0008046359871514142
step: 70, loss: 0.0003421236469876021
step: 80, loss: 0.006103364750742912
step: 90, loss: 0.05046052113175392
step: 100, loss: 0.007223143707960844
step: 110, loss: 0.00044656250975094736
step: 120, loss: 0.0407801978290081
step: 130, loss: 0.061839353293180466
step: 140, loss: 0.0003093378327321261
step: 150, loss: 0.0001362603943562135
step: 160, loss: 0.014822503551840782
step: 170, loss: 0.020406784489750862
step: 180, loss: 0.0005210363888181746
step: 190, loss: 0.00043079874012619257

```

```

step: 200, loss: 0.0030291611328721046
step: 210, loss: 0.0003159455372951925
step: 220, loss: 0.0005044181016273797
step: 230, loss: 0.0004644641012419015
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.9653631284916202, f1=0.9634551495016611,
best_f1=0.9698324022346367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005675662541761994
step: 10, loss: 0.0003686398558784276
step: 20, loss: 0.0005065240547992289
step: 30, loss: 0.008548957295715809
step: 40, loss: 0.0036500790156424046
step: 50, loss: 0.010914798825979233
step: 60, loss: 0.021523769944906235
step: 70, loss: 0.002907311776652932
step: 80, loss: 0.005823428742587566
step: 90, loss: 0.0006962757906876504
step: 100, loss: 0.0005460020038299263
step: 110, loss: 0.0297967791557312
step: 120, loss: 0.00068213528720662
step: 130, loss: 0.009464799426496029
step: 140, loss: 0.0003352975472807884
step: 150, loss: 0.00034257915103808045
step: 160, loss: 0.024725519120693207
step: 170, loss: 0.00016324750322382897
step: 180, loss: 0.0008405277039855719
step: 190, loss: 0.0018555474234744906
step: 200, loss: 0.09361326694488525
step: 210, loss: 0.0018741564126685262
step: 220, loss: 0.0005355238681659102
step: 230, loss: 0.0002833804173860699
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.963718820861678, f1=0.9686800894854586,
best_f1=0.9698324022346367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07315387576818466
step: 10, loss: 0.0005749044357798994
step: 20, loss: 0.00017785483214538544
step: 30, loss: 0.0006181808421388268

```



```

step: 40, loss: 0.005137530621141195
step: 50, loss: 0.0005419256631284952
step: 60, loss: 0.081812784075737
step: 70, loss: 0.0006066395435482264
step: 80, loss: 0.00016307947225868702
step: 90, loss: 0.004121021833270788
step: 100, loss: 0.0012588928220793605
step: 110, loss: 0.005087791010737419
step: 120, loss: 0.0026229405775666237
step: 130, loss: 0.0012774395290762186
step: 140, loss: 0.00016472079732920974
step: 150, loss: 0.14372308552265167
step: 160, loss: 0.000350294605595991
step: 170, loss: 0.01120296772569418
step: 180, loss: 0.00048033983330242336
step: 190, loss: 0.005021101329475641
step: 200, loss: 0.0030645192600786686
step: 210, loss: 0.00046959100291132927
step: 220, loss: 0.0014449061127379537
step: 230, loss: 0.023531073704361916
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.967525195968645, f1=0.96353591160221, best_f1=0.96353591160221
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017646126449108124
step: 10, loss: 0.006356842815876007
step: 20, loss: 0.0003411652578506619
step: 30, loss: 0.00042411984759382904
step: 40, loss: 0.00018684676615521312
step: 50, loss: 0.0026563459541648626
step: 60, loss: 0.00022528663976117969
step: 70, loss: 0.0006706694257445633
step: 80, loss: 0.00018259490025229752
step: 90, loss: 0.00019614827760960907
step: 100, loss: 0.00028077609022147954
step: 110, loss: 0.0002962453290820122
step: 120, loss: 0.00010154205665457994
step: 130, loss: 0.0020802749786525965
step: 140, loss: 0.0002335321478312835
step: 150, loss: 0.0001400929322699085
step: 160, loss: 0.0003406466857995838

```

```

step: 170, loss: 0.006880174856632948
step: 180, loss: 0.0019506571115925908
step: 190, loss: 0.00102844403591007
step: 200, loss: 0.001405910006724298
step: 210, loss: 0.00022035832807887346
step: 220, loss: 0.037151627242565155
step: 230, loss: 0.0391756147146225
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9599999999999999, f1=0.9588431590656283,
best_f1=0.96353591160221
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015426719619426876
step: 10, loss: 0.00149127549957484
step: 20, loss: 0.0038209501653909683
step: 30, loss: 0.001421432476490736
step: 40, loss: 0.00014866287529002875
step: 50, loss: 0.004249874502420425
step: 60, loss: 0.0005770345451310277
step: 70, loss: 0.000512156926561147
step: 80, loss: 0.09548552334308624
step: 90, loss: 0.018968798220157623
step: 100, loss: 0.0003039421862922609
step: 110, loss: 0.005232402589172125
step: 120, loss: 0.039329033344984055
step: 130, loss: 0.00015556503785774112
step: 140, loss: 0.07574116438627243
step: 150, loss: 0.00012355616490822285
step: 160, loss: 0.0036524091847240925
step: 170, loss: 0.00013255106750875711
step: 180, loss: 0.001419505919329822
step: 190, loss: 0.0003100506728515029
step: 200, loss: 0.0002623398322612047
step: 210, loss: 0.0007539559155702591
step: 220, loss: 0.03312515839934349
step: 230, loss: 0.015228857286274433
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.9696287964004499, f1=0.9644444444444443,
best_f1=0.9644444444444443

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009589939727447927
step: 10, loss: 7.132847531465814e-05
step: 20, loss: 0.0001196901430375874
step: 30, loss: 0.00030699738999828696
step: 40, loss: 6.973068229854107e-05
step: 50, loss: 0.02838090993463993
step: 60, loss: 0.030825292691588402
step: 70, loss: 0.0008553875959478319
step: 80, loss: 0.034518834203481674
step: 90, loss: 0.0002010303142014891
step: 100, loss: 0.00034390779910609126
step: 110, loss: 0.0007403485360555351
step: 120, loss: 0.015875767916440964
step: 130, loss: 0.000657634052913636
step: 140, loss: 0.027697589248418808
step: 150, loss: 0.0001267455518245697
step: 160, loss: 0.0013202237896621227
step: 170, loss: 0.011047097854316235
step: 180, loss: 0.001828051288612187
step: 190, loss: 0.00023232644889503717
step: 200, loss: 9.08023357624188e-05
step: 210, loss: 0.00028496055165305734
step: 220, loss: 0.001066923956386745
step: 230, loss: 0.09489426016807556
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9678135405105438, f1=0.9643652561247216,
best_f1=0.9644444444444443
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010413239942863584
step: 10, loss: 0.00025446509243920445
step: 20, loss: 5.2871444495394826e-05
step: 30, loss: 0.00039919003029353917
step: 40, loss: 0.014532004483044147
step: 50, loss: 0.001604909310117364
step: 60, loss: 0.00024243337975349277
step: 70, loss: 0.0004948821733705699
step: 80, loss: 0.0018202911596745253
step: 90, loss: 0.00020222218881826848
step: 100, loss: 0.0006291804602369666
step: 110, loss: 0.002855700673535466
step: 120, loss: 0.00029223793535493314

```

```

step: 130, loss: 6.374313670676202e-05
step: 140, loss: 0.00011757940956158563
step: 150, loss: 8.312152203870937e-05
step: 160, loss: 0.0001991699100472033
step: 170, loss: 0.011815246194601059
step: 180, loss: 0.0003739074745681137
step: 190, loss: 0.0002700352342799306
step: 200, loss: 0.00014449992158915848
step: 210, loss: 7.329422805923969e-05
step: 220, loss: 7.258024561451748e-05
step: 230, loss: 0.01725686714053154
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9650507328072153, f1=0.9598214285714285,
best_f1=0.9644444444444443
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00046884402399882674
step: 10, loss: 0.0031038119923323393
step: 20, loss: 5.784104723716155e-05
step: 30, loss: 0.0056206537410616875
step: 40, loss: 5.576628973358311e-05
step: 50, loss: 8.024970884434879e-05
step: 60, loss: 0.017400875687599182
step: 70, loss: 7.585166895296425e-05
step: 80, loss: 7.802416803315282e-05
step: 90, loss: 0.0005542318103834987
step: 100, loss: 4.823138806386851e-05
step: 110, loss: 0.00015450552746187896
step: 120, loss: 9.316421346738935e-05
step: 130, loss: 0.000258623156696558
step: 140, loss: 8.627025090390816e-05
step: 150, loss: 6.651013245573267e-05
step: 160, loss: 0.014583217911422253
step: 170, loss: 8.855303894961253e-05
step: 180, loss: 0.00023683109611738473
step: 190, loss: 0.0001595442881807685
step: 200, loss: 0.009873870760202408
step: 210, loss: 6.916301936144009e-05
step: 220, loss: 0.020203853026032448
step: 230, loss: 0.0001389398385072127
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9684684684684683, f1=0.9644444444444443,
best_f1=0.9644444444444443

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005173937533982098
step: 10, loss: 0.00013038313772995025
step: 20, loss: 0.00014265165373217314
step: 30, loss: 0.016853565350174904
step: 40, loss: 0.00016991225129459053
step: 50, loss: 8.715103467693552e-05
step: 60, loss: 0.00012944494665134698
step: 70, loss: 0.0001936344342539087
step: 80, loss: 0.00017310694965999573
step: 90, loss: 0.0008677459554746747
step: 100, loss: 0.00040976787568069994
step: 110, loss: 0.00014354677114170045
step: 120, loss: 0.0009034270769916475
step: 130, loss: 0.00015643879305571318
step: 140, loss: 0.005492402706295252
step: 150, loss: 0.031218362972140312
step: 160, loss: 5.1442777476040646e-05
step: 170, loss: 0.0006013956153765321
step: 180, loss: 0.00017557726823724806
step: 190, loss: 7.387147343251854e-05
step: 200, loss: 0.00024098405265249312
step: 210, loss: 6.577120075235143e-05
step: 220, loss: 0.0003735897480510175
step: 230, loss: 0.0003259656368754804
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9694224235560589, f1=0.9608938547486034,
best_f1=0.9644444444444443
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.173134195501916e-05
step: 10, loss: 5.174945545149967e-05
step: 20, loss: 0.001141634420491755
step: 30, loss: 0.0002166702033719048
step: 40, loss: 0.00022228265879675746
step: 50, loss: 0.00016059218614827842
step: 60, loss: 5.687322118319571e-05
step: 70, loss: 6.002910595270805e-05
step: 80, loss: 7.542177627328783e-05
step: 90, loss: 0.00020908736041747034
step: 100, loss: 0.0054697077721357346
step: 110, loss: 0.00030308758141472936
step: 120, loss: 4.649172115023248e-05

```

```

step: 130, loss: 0.00011030451423721388
step: 140, loss: 4.559722219710238e-05
step: 150, loss: 8.054055069806054e-05
step: 160, loss: 0.0001924390671774745
step: 170, loss: 4.8797050112625584e-05
step: 180, loss: 7.321698649320751e-05
step: 190, loss: 0.0018140054307878017
step: 200, loss: 6.368012691382319e-05
step: 210, loss: 0.02776934765279293
step: 220, loss: 0.00012162708299001679
step: 230, loss: 3.098577508353628e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9694224235560589, f1=0.9620535714285715,
best_f1=0.9644444444444443
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.7720146792707965e-05
step: 10, loss: 0.0001809443929232657
step: 20, loss: 9.408619371242821e-05
step: 30, loss: 0.0003695484483614564
step: 40, loss: 6.815903907408938e-05
step: 50, loss: 7.668664329685271e-05
step: 60, loss: 4.524370524450205e-05
step: 70, loss: 0.0003371161874383688
step: 80, loss: 0.013434135355055332
step: 90, loss: 3.0490311473840848e-05
step: 100, loss: 5.338424307410605e-05
step: 110, loss: 5.755278471042402e-05
step: 120, loss: 7.490580901503563e-05
step: 130, loss: 0.00016965257236734033
step: 140, loss: 0.00014944533177185804
step: 150, loss: 0.0003610651765484363
step: 160, loss: 8.238332520704716e-05
step: 170, loss: 3.9195634599309415e-05
step: 180, loss: 5.140220309840515e-05
step: 190, loss: 0.00042360328370705247
step: 200, loss: 4.944550892105326e-05
step: 210, loss: 0.03746093809604645
step: 220, loss: 6.325967842712998e-05
step: 230, loss: 7.10302556399256e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or

```

```

loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 15: dev_f1=0.9705215419501134, f1=0.9609810479375697,
best_f1=0.9609810479375697
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2473it [00:07, 312.57it/s]
load_f1 = 0.9705882352941176
real_f1 = 0.9694915254237287
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:11, 384.71it/s]

```

### 1.3.2 DBLP-GoogleScholar - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Dirty/DBLP-GoogleScholar \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm distilbert \
  --fp16 \
  --save_model \
  --da span_shuffle \
  --da attr_del \
  --summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-GoogleScholar/.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Dirty/DBLP-GoogleScholar \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm distilbert \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/
```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```
enabled          : True
```



```

opt_level                : 02
cast_model_type          : torch.float16
patch_torch_functions    : False
keep_batchnorm_fp32     : True
master_weights           : True
loss_scale               : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.7968710660934448
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4588870704174042
step: 20, loss: 0.4950137138366699
step: 30, loss: 0.4503419101238251
step: 40, loss: 0.3774471879005432
step: 50, loss: 0.24403983354568481
step: 60, loss: 0.23045320808887482
step: 70, loss: 0.10644803941249847
step: 80, loss: 0.1431143432855606
step: 90, loss: 0.21971310675144196
step: 100, loss: 0.2580089271068573
step: 110, loss: 0.09373287856578827
step: 120, loss: 0.06358713656663895
step: 130, loss: 0.019797753542661667
step: 140, loss: 0.12465174496173859
step: 150, loss: 0.06523783504962921
step: 160, loss: 0.18320529162883759
step: 170, loss: 0.1275460422039032
step: 180, loss: 0.10922011733055115
step: 190, loss: 0.05544634535908699
step: 200, loss: 0.08862265944480896
step: 210, loss: 0.08851314336061478
step: 220, loss: 0.1663406938314438
step: 230, loss: 0.07963055372238159
step: 240, loss: 0.09995976090431213

```

```

step: 250, loss: 0.05870873108506203
step: 260, loss: 0.04448387399315834
step: 270, loss: 0.016654042527079582
step: 280, loss: 0.12467319518327713
step: 290, loss: 0.14918754994869232
step: 300, loss: 0.09787188470363617
step: 310, loss: 0.15620645880699158
step: 320, loss: 0.12195289134979248
step: 330, loss: 0.17440786957740784
step: 340, loss: 0.2616775929927826
step: 350, loss: 0.11146049201488495
step: 360, loss: 0.11710640788078308
step: 370, loss: 0.1638319194316864
step: 380, loss: 0.2445455640554428
step: 390, loss: 0.13730531930923462
step: 400, loss: 0.11225768178701401
step: 410, loss: 0.03799201175570488
step: 420, loss: 0.044336557388305664
step: 430, loss: 0.07289464771747589
step: 440, loss: 0.18177518248558044
step: 450, loss: 0.06532780081033707
step: 460, loss: 0.07051662355661392
step: 470, loss: 0.23113079369068146
step: 480, loss: 0.25238561630249023
step: 490, loss: 0.07898245751857758
step: 500, loss: 0.01570316031575203
step: 510, loss: 0.10068803280591965
step: 520, loss: 0.06571657210588455
step: 530, loss: 0.06104730814695358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9180778032036614, f1=0.9054545454545454,
best_f1=0.9054545454545454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14440569281578064
step: 10, loss: 0.16022919118404388
step: 20, loss: 0.1497599184513092
step: 30, loss: 0.045275501906871796
step: 40, loss: 0.00340961292386055
step: 50, loss: 0.15493226051330566
step: 60, loss: 0.13507500290870667

```

step: 70, loss: 0.11757250875234604  
step: 80, loss: 0.023841122165322304  
step: 90, loss: 0.003354305401444435  
step: 100, loss: 0.2572605609893799  
step: 110, loss: 0.053534138947725296  
step: 120, loss: 0.10418549925088882  
step: 130, loss: 0.007070058491080999  
step: 140, loss: 0.02126777544617653  
step: 150, loss: 0.07543567568063736  
step: 160, loss: 0.07396256923675537  
step: 170, loss: 0.12194329500198364  
step: 180, loss: 0.20708918571472168  
step: 190, loss: 0.025080619379878044  
step: 200, loss: 0.06907664984464645  
step: 210, loss: 0.013634948059916496  
step: 220, loss: 0.21254615485668182  
step: 230, loss: 0.04207345098257065  
step: 240, loss: 0.18076977133750916  
step: 250, loss: 0.030159318819642067  
step: 260, loss: 0.012684019282460213  
step: 270, loss: 0.13953834772109985  
step: 280, loss: 0.24265286326408386  
step: 290, loss: 0.1477510929107666  
step: 300, loss: 0.039949435740709305  
step: 310, loss: 0.03727808967232704  
step: 320, loss: 0.06930028647184372  
step: 330, loss: 0.010844786651432514  
step: 340, loss: 0.004611378535628319  
step: 350, loss: 0.07895055413246155  
step: 360, loss: 0.061138033866882324  
step: 370, loss: 0.016693921759724617  
step: 380, loss: 0.06869655102491379  
step: 390, loss: 0.02088005840778351  
step: 400, loss: 0.07154905796051025  
step: 410, loss: 0.0007805959903635085  
step: 420, loss: 0.02494588866829872  
step: 430, loss: 0.02670213393867016  
step: 440, loss: 0.014606758020818233  
step: 450, loss: 0.06943190842866898  
step: 460, loss: 0.27969232201576233  
step: 470, loss: 0.023152055218815804  
step: 480, loss: 0.20756803452968597  
step: 490, loss: 0.07243867218494415  
step: 500, loss: 0.014926942065358162  
step: 510, loss: 0.04092821478843689  
step: 520, loss: 0.053430281579494476  
step: 530, loss: 0.14842933416366577  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning:

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.924791086350975, f1=0.9194974406700791,
best_f1=0.9194974406700791
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017080560326576233
step: 10, loss: 0.04597119614481926
step: 20, loss: 0.1837920844554901
step: 30, loss: 0.13700748980045319
step: 40, loss: 0.008013888262212276
step: 50, loss: 0.10206802189350128
step: 60, loss: 0.049499642103910446
step: 70, loss: 0.010242489166557789
step: 80, loss: 0.04586559906601906
step: 90, loss: 0.3052269220352173
step: 100, loss: 0.04932268336415291
step: 110, loss: 0.03181460127234459
step: 120, loss: 0.00691812951117754
step: 130, loss: 0.027202554047107697
step: 140, loss: 0.03874516114592552
step: 150, loss: 0.014586756937205791
step: 160, loss: 0.02261887863278389
step: 170, loss: 0.010436922311782837
step: 180, loss: 0.12053722888231277
step: 190, loss: 0.04882329702377319
step: 200, loss: 0.062045298516750336
step: 210, loss: 0.19014038145542145
step: 220, loss: 0.042616426944732666
step: 230, loss: 0.018157538026571274
step: 240, loss: 0.016451917588710785
step: 250, loss: 0.09369798004627228
step: 260, loss: 0.0041721286252141
step: 270, loss: 0.005034558475017548
step: 280, loss: 0.01435667835175991
step: 290, loss: 0.06559331715106964
step: 300, loss: 0.1034892201423645
step: 310, loss: 0.1710788756608963
step: 320, loss: 0.20851252973079681
step: 330, loss: 0.06403707712888718
step: 340, loss: 0.010415386408567429
step: 350, loss: 0.04147755727171898
step: 360, loss: 0.033620838075876236

```

```

step: 370, loss: 0.007718503009527922
step: 380, loss: 0.032343991100788116
step: 390, loss: 0.08598022162914276
step: 400, loss: 0.02141747996211052
step: 410, loss: 0.010224839672446251
step: 420, loss: 0.00811222568154335
step: 430, loss: 0.02190639264881611
step: 440, loss: 0.04931307211518288
step: 450, loss: 0.2925710082054138
step: 460, loss: 0.15003997087478638
step: 470, loss: 0.02289193868637085
step: 480, loss: 0.018083155155181885
step: 490, loss: 0.016996312886476517
step: 500, loss: 0.052097514271736145
step: 510, loss: 0.003687243675813079
step: 520, loss: 0.004360634833574295
step: 530, loss: 0.07988206297159195
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.9204705882352942, f1=0.9135338345864661,
best_f1=0.9194974406700791
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0480872206389904
step: 10, loss: 0.027001116424798965
step: 20, loss: 0.033773064613342285
step: 30, loss: 0.007742106914520264
step: 40, loss: 0.02338336408138275
step: 50, loss: 0.05568699538707733
step: 60, loss: 0.0025076745077967644
step: 70, loss: 0.007510061375796795
step: 80, loss: 0.03678664192557335
step: 90, loss: 0.024223508313298225
step: 100, loss: 0.010810857638716698
step: 110, loss: 0.005134479142725468
step: 120, loss: 0.004149157553911209
step: 130, loss: 0.1689797043800354
step: 140, loss: 0.17978791892528534
step: 150, loss: 0.003910842351615429
step: 160, loss: 0.01530400663614273
step: 170, loss: 0.008189858868718147
step: 180, loss: 0.0047505199909210205
step: 190, loss: 0.0036460035480558872
step: 200, loss: 0.021681496873497963
step: 210, loss: 0.1184917762875557
step: 220, loss: 0.003937935922294855

```

```

step: 230, loss: 0.24500754475593567
step: 240, loss: 0.046401120722293854
step: 250, loss: 0.004564111120998859
step: 260, loss: 0.13709957897663116
step: 270, loss: 0.03380673751235008
step: 280, loss: 0.01143616158515215
step: 290, loss: 0.02786763198673725
step: 300, loss: 0.00181258050724864
step: 310, loss: 0.0043691666796803474
step: 320, loss: 0.02208196371793747
step: 330, loss: 0.06867720931768417
step: 340, loss: 0.0074424901977181435
step: 350, loss: 0.04599442705512047
step: 360, loss: 0.009286811575293541
step: 370, loss: 0.05784761160612106
step: 380, loss: 0.05284601449966431
step: 390, loss: 0.010754209943115711
step: 400, loss: 0.024251727387309074
step: 410, loss: 0.015141644515097141
step: 420, loss: 0.005372214131057262
step: 430, loss: 0.019634734839200974
step: 440, loss: 0.09753914177417755
step: 450, loss: 0.0025468426756560802
step: 460, loss: 0.08571850508451462
step: 470, loss: 0.07824650406837463
step: 480, loss: 0.021880848333239555
step: 490, loss: 0.021337278187274933
step: 500, loss: 0.04399140924215317
step: 510, loss: 0.10076408088207245
step: 520, loss: 0.048287659883499146
step: 530, loss: 0.0027133189141750336
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.9194974406700791, f1=0.9215686274509803,
best_f1=0.9194974406700791
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00495497602969408
step: 10, loss: 0.03866606205701828
step: 20, loss: 0.04490024968981743
step: 30, loss: 0.002839642111212015
step: 40, loss: 0.07228822261095047
step: 50, loss: 0.006777304690331221
step: 60, loss: 0.0027236584573984146
step: 70, loss: 0.0036507423501461744
step: 80, loss: 0.002253949409350753

```

```
step: 90, loss: 0.021244298666715622
step: 100, loss: 0.0032934413757175207
step: 110, loss: 0.0037328817415982485
step: 120, loss: 0.007641106843948364
step: 130, loss: 0.0017437310889363289
step: 140, loss: 0.0011615806724876165
step: 150, loss: 0.0014986679889261723
step: 160, loss: 0.014266704209148884
step: 170, loss: 0.02757100947201252
step: 180, loss: 0.004594959318637848
step: 190, loss: 0.0023625504691153765
step: 200, loss: 0.07341853529214859
step: 210, loss: 0.0039206393994390965
step: 220, loss: 0.0011016130447387695
step: 230, loss: 0.001268188003450632
step: 240, loss: 0.07925165444612503
step: 250, loss: 0.0005895032081753016
step: 260, loss: 0.07728973031044006
step: 270, loss: 0.06442943960428238
step: 280, loss: 0.08410529792308807
step: 290, loss: 0.0041730026714503765
step: 300, loss: 0.04034439101815224
step: 310, loss: 0.0011988197220489383
step: 320, loss: 0.00976533442735672
step: 330, loss: 0.004687233362346888
step: 340, loss: 0.0063133155927062035
step: 350, loss: 0.0021040274295955896
step: 360, loss: 0.00483450572937727
step: 370, loss: 0.07562313973903656
step: 380, loss: 0.028535256162285805
step: 390, loss: 0.0029690305236727
step: 400, loss: 0.024769067764282227
step: 410, loss: 0.0015451166545972228
step: 420, loss: 0.0006120260222814977
step: 430, loss: 0.11985430121421814
step: 440, loss: 0.03276451304554939
step: 450, loss: 0.03271351754665375
step: 460, loss: 0.018214868381619453
step: 470, loss: 0.023648984730243683
step: 480, loss: 0.06011735647916794
step: 490, loss: 0.032602109014987946
step: 500, loss: 0.01501951739192009
step: 510, loss: 0.039518196135759354
step: 520, loss: 0.007178163155913353
step: 530, loss: 0.0014556117821484804
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.9255813953488372, f1=0.9199063231850116,
best_f1=0.9199063231850116
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003970821853727102
step: 10, loss: 0.0016489119734615088
step: 20, loss: 0.000899057777132839
step: 30, loss: 0.009490127675235271
step: 40, loss: 0.20212194323539734
step: 50, loss: 0.005952935200184584
step: 60, loss: 0.014977176673710346
step: 70, loss: 0.06353705376386642
step: 80, loss: 0.003995611798018217
step: 90, loss: 0.009032521396875381
step: 100, loss: 0.012023475021123886
step: 110, loss: 0.014102919027209282
step: 120, loss: 0.002101857913658023
step: 130, loss: 0.0027261406648904085
step: 140, loss: 0.001600343268364668
step: 150, loss: 0.0017793620936572552
step: 160, loss: 0.001262778416275978
step: 170, loss: 0.0022882786579430103
step: 180, loss: 0.0010566527489572763
step: 190, loss: 0.06722662597894669
step: 200, loss: 0.008596522733569145
step: 210, loss: 0.0071508316323161125
step: 220, loss: 0.0021049864590168
step: 230, loss: 0.0005799625068902969
step: 240, loss: 0.0033524527680128813
step: 250, loss: 0.004719584248960018
step: 260, loss: 0.05682859942317009
step: 270, loss: 0.001353534171357751
step: 280, loss: 0.009336578659713268
step: 290, loss: 0.0009500641608610749
step: 300, loss: 0.061124395579099655
step: 310, loss: 0.014615335501730442
step: 320, loss: 0.08497367799282074
step: 330, loss: 0.03302565589547157
step: 340, loss: 0.06084572151303291
step: 350, loss: 0.10161957144737244
step: 360, loss: 0.0007496022153645754
step: 370, loss: 0.009300353936851025
step: 380, loss: 0.0017553700599819422

```



```

step: 390, loss: 0.06110750883817673
step: 400, loss: 0.00024327145365532488
step: 410, loss: 0.0015344069106504321
step: 420, loss: 0.012316945940256119
step: 430, loss: 0.00033914032974280417
step: 440, loss: 0.028810618445277214
step: 450, loss: 0.0012438183184713125
step: 460, loss: 0.049400463700294495
step: 470, loss: 0.0036351229064166546
step: 480, loss: 0.012661690823733807
step: 490, loss: 0.0011598992859944701
step: 500, loss: 0.001875304733403027
step: 510, loss: 0.003271284978836775
step: 520, loss: 0.0027989803347736597
step: 530, loss: 0.011317485012114048
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9227188081936686, f1=0.9131047440112728,
best_f1=0.9199063231850116
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003220595186576247
step: 10, loss: 0.01269399281591177
step: 20, loss: 0.003572812071070075
step: 30, loss: 0.0005113806109875441
step: 40, loss: 0.00024172771372832358
step: 50, loss: 0.0509042963385582
step: 60, loss: 0.0023987474851310253
step: 70, loss: 0.0007602550904266536
step: 80, loss: 0.007916565053164959
step: 90, loss: 0.002730804029852152
step: 100, loss: 0.0016444725915789604
step: 110, loss: 0.046850696206092834
step: 120, loss: 0.011517866514623165
step: 130, loss: 0.00013096368638798594
step: 140, loss: 0.04264051839709282
step: 150, loss: 9.402076830156147e-05
step: 160, loss: 0.00024843058781698346
step: 170, loss: 0.0007075549219734967
step: 180, loss: 0.00220381747931242
step: 190, loss: 9.950977982953191e-05
step: 200, loss: 0.0010851842816919088
step: 210, loss: 0.002157118869945407
step: 220, loss: 0.0011697645531967282
step: 230, loss: 0.0008288713870570064
step: 240, loss: 0.0025148813147097826

```

```

step: 250, loss: 0.0012427627807483077
step: 260, loss: 0.0315825417637825
step: 270, loss: 0.0010096896439790726
step: 280, loss: 0.0010336796985939145
step: 290, loss: 0.00341031770221889
step: 300, loss: 0.0009637378971092403
step: 310, loss: 0.005639077164232731
step: 320, loss: 0.030819270759820938
step: 330, loss: 0.0001491253642598167
step: 340, loss: 0.003425776259973645
step: 350, loss: 0.009198176674544811
step: 360, loss: 0.0020135235972702503
step: 370, loss: 0.00040652614552527666
step: 380, loss: 0.001862534205429256
step: 390, loss: 0.0008132763905450702
step: 400, loss: 0.01253652386367321
step: 410, loss: 0.020473994314670563
step: 420, loss: 0.0002331286232220009
step: 430, loss: 0.0001088419885491021
step: 440, loss: 0.00026076383073814213
step: 450, loss: 0.016330542042851448
step: 460, loss: 0.0019660175312310457
step: 470, loss: 0.06391721963882446
step: 480, loss: 0.0014942977577447891
step: 490, loss: 0.008428863249719143
step: 500, loss: 0.00037037988658994436
step: 510, loss: 0.005842606537044048
step: 520, loss: 0.004009009804576635
step: 530, loss: 0.000330063485307619
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9238770685579196, f1=0.9227144203581528,
best_f1=0.9199063231850116
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0037637269124388695
step: 10, loss: 0.041537199169397354
step: 20, loss: 0.02142467349767685
step: 30, loss: 0.004242236260324717
step: 40, loss: 0.07341492176055908
step: 50, loss: 0.004248740151524544
step: 60, loss: 0.00533508462831378
step: 70, loss: 0.0009395344532094896
step: 80, loss: 8.640066516818479e-05
step: 90, loss: 0.001970387063920498
step: 100, loss: 0.003748913761228323

```

```

step: 110, loss: 0.0008677869336679578
step: 120, loss: 0.004803983960300684
step: 130, loss: 0.0009431624785065651
step: 140, loss: 5.614127803710289e-05
step: 150, loss: 0.037064939737319946
step: 160, loss: 0.0004001337510999292
step: 170, loss: 0.03265928849577904
step: 180, loss: 0.00217783753760159
step: 190, loss: 0.003931316081434488
step: 200, loss: 0.0004956712364219129
step: 210, loss: 0.06350895762443542
step: 220, loss: 0.0016012485139071941
step: 230, loss: 0.0021683587692677975
step: 240, loss: 0.005366852972656488
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 250, loss: 0.058211907744407654
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 260, loss: 0.2261921614408493
step: 270, loss: 5.072136264061555e-05
step: 280, loss: 0.001295543392188847
step: 290, loss: 0.0022225764114409685
step: 300, loss: 0.00021871265198569745
step: 310, loss: 0.043644070625305176
step: 320, loss: 0.006749613210558891
step: 330, loss: 0.005754007026553154
step: 340, loss: 0.0013892214046791196
step: 350, loss: 0.02431369572877884
step: 360, loss: 0.0013782165478914976
step: 370, loss: 0.0018392017809674144
step: 380, loss: 0.05499658361077309
step: 390, loss: 0.0007696514367125928
step: 400, loss: 0.08987249433994293
step: 410, loss: 0.0006222177180461586
step: 420, loss: 0.0007050951826386154
step: 430, loss: 0.000526034040376544
step: 440, loss: 0.0382964126765728
step: 450, loss: 0.0014854957116767764
step: 460, loss: 0.00019528827397152781
step: 470, loss: 0.0007395996944978833
step: 480, loss: 0.00027093931566923857
step: 490, loss: 0.00016738724661991
step: 500, loss: 0.00016448902897536755
step: 510, loss: 9.407660400029272e-05
step: 520, loss: 0.000608751317486167
step: 530, loss: 0.0002828461001627147
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```
epoch 8: dev_f1=0.9242700729927007, f1=0.9206349206349206,  
best_f1=0.9199063231850116  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.004281553439795971  
step: 10, loss: 0.0006742752739228308  
step: 20, loss: 0.00012262801465112716  
step: 30, loss: 0.0003972419654019177  
step: 40, loss: 0.0008624355541542172  
step: 50, loss: 0.0002589826472103596  
step: 60, loss: 0.08145929872989655  
step: 70, loss: 0.08599816262722015  
step: 80, loss: 0.0024214559234678745  
step: 90, loss: 0.00010678747639758512  
step: 100, loss: 0.0003388224868103862  
step: 110, loss: 0.001047119265422225  
step: 120, loss: 0.0006629563868045807  
step: 130, loss: 0.00014844507677480578  
step: 140, loss: 0.005104833748191595  
step: 150, loss: 0.0001676736428635195  
step: 160, loss: 0.0035568864550441504  
step: 170, loss: 0.00044662749860435724  
step: 180, loss: 0.001139301573857665  
step: 190, loss: 0.0057940236292779446  
step: 200, loss: 0.0008807548438198864  
step: 210, loss: 0.0005287131643854082  
step: 220, loss: 0.016601286828517914  
step: 230, loss: 0.0009700777009129524  
step: 240, loss: 0.0024509804788976908  
step: 250, loss: 0.002731184009462595  
step: 260, loss: 0.0030732937157154083  
step: 270, loss: 0.001421290566213429  
step: 280, loss: 0.00012684709508903325  
step: 290, loss: 5.062992931925692e-05  
step: 300, loss: 0.00023659842554479837  
step: 310, loss: 0.0013828836381435394  
step: 320, loss: 0.001835510483942926  
step: 330, loss: 0.0009634580928832293  
step: 340, loss: 6.790355109842494e-05  
step: 350, loss: 0.0012177779572084546  
step: 360, loss: 0.004202776588499546  
step: 370, loss: 0.000657636730466038  
step: 380, loss: 0.0006699677905999124  
step: 390, loss: 0.00015768909361213446  
step: 400, loss: 0.00977308489382267  
step: 410, loss: 0.0001465282985009253  
step: 420, loss: 0.001640453701838851
```

```
step: 430, loss: 8.104000153252855e-05
step: 440, loss: 0.0004634308279491961
step: 450, loss: 0.0001963700051419437
step: 460, loss: 0.006538308225572109
step: 470, loss: 3.8171860069269314e-05
step: 480, loss: 0.0006507051875814795
step: 490, loss: 0.0021412840578705072
step: 500, loss: 0.0034594370517879725
step: 510, loss: 0.018596109002828598
step: 520, loss: 0.000386189145501703
step: 530, loss: 8.351392898475751e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9167816091954023, f1=0.9217391304347826,
best_f1=0.9199063231850116
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016663186252117157
step: 10, loss: 3.293455665698275e-05
step: 20, loss: 0.003489609807729721
step: 30, loss: 0.0005151015357114375
step: 40, loss: 7.688150071771815e-05
step: 50, loss: 0.003613634267821908
step: 60, loss: 8.234671986429021e-05
step: 70, loss: 0.0003146928793285042
step: 80, loss: 0.000881813233718276
step: 90, loss: 0.000615648808889091
step: 100, loss: 0.00016737569239921868
step: 110, loss: 0.004335158970206976
step: 120, loss: 0.00036726443795487285
step: 130, loss: 0.00042253112769685686
step: 140, loss: 0.029366284608840942
step: 150, loss: 0.004466644488275051
step: 160, loss: 0.004866814240813255
step: 170, loss: 0.00012583153147716075
step: 180, loss: 0.003485990222543478
step: 190, loss: 0.008767144754529
step: 200, loss: 0.0037974058650434017
step: 210, loss: 0.0034311439376324415
step: 220, loss: 0.0005103647708892822
step: 230, loss: 0.0017496615182608366
step: 240, loss: 8.269018871942535e-05
step: 250, loss: 0.022820554673671722
step: 260, loss: 0.09962131083011627
step: 270, loss: 0.00016080689965747297
step: 280, loss: 4.140082819503732e-05
```

```

step: 290, loss: 0.0006434953538700938
step: 300, loss: 0.014938535168766975
step: 310, loss: 0.0006412248476408422
step: 320, loss: 0.00011367943079676479
step: 330, loss: 0.0003597991308197379
step: 340, loss: 0.0003374749794602394
step: 350, loss: 0.013219532556831837
step: 360, loss: 9.557753946864977e-05
step: 370, loss: 0.0090740155428648
step: 380, loss: 0.0033730841241776943
step: 390, loss: 0.00012386094022076577
step: 400, loss: 0.0003761453554034233
step: 410, loss: 0.0001068356359610334
step: 420, loss: 0.01913950964808464
step: 430, loss: 0.0002745790989138186
step: 440, loss: 0.00027841885457746685
step: 450, loss: 0.0155861284583807
step: 460, loss: 0.06504224240779877
step: 470, loss: 0.0012080835876986384
step: 480, loss: 0.0004135374038014561
step: 490, loss: 0.1560916304588318
step: 500, loss: 0.046509936451911926
step: 510, loss: 0.0053770761005580425
step: 520, loss: 0.005660895258188248
step: 530, loss: 0.0028966290410608053
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9191685912240185, f1=0.9225092250922509,
best_f1=0.9199063231850116
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003117230022326112
step: 10, loss: 0.0925212949514389
step: 20, loss: 0.00012776943913195282
step: 30, loss: 0.00048570468788966537
step: 40, loss: 0.0014720456674695015
step: 50, loss: 0.0011358411284163594
step: 60, loss: 0.0036978833377361298
step: 70, loss: 0.00016702336142770946
step: 80, loss: 0.0009247945272363722
step: 90, loss: 0.001118781277909875
step: 100, loss: 0.00014443896361626685
step: 110, loss: 0.0008134585805237293
step: 120, loss: 0.001477244310081005
step: 130, loss: 0.11342670768499374
step: 140, loss: 0.00026472180616110563

```

```

step: 150, loss: 0.00291884015314281
step: 160, loss: 0.009063002653419971
step: 170, loss: 0.00184713548514992
step: 180, loss: 0.0005785381072200835
step: 190, loss: 0.0009192110155709088
step: 200, loss: 0.00991327129304409
step: 210, loss: 0.0047583384439349174
step: 220, loss: 0.0003773628268390894
step: 230, loss: 0.00903834868222475
step: 240, loss: 0.00010680006380425766
step: 250, loss: 0.0012312184553593397
step: 260, loss: 0.0003869817010127008
step: 270, loss: 0.004493284039199352
step: 280, loss: 0.0040142429061234
step: 290, loss: 0.0026252816896885633
step: 300, loss: 0.021611254662275314
step: 310, loss: 0.0070971157401800156
step: 320, loss: 0.019268738105893135
step: 330, loss: 0.01914423145353794
step: 340, loss: 0.004040293395519257
step: 350, loss: 0.0003790615010075271
step: 360, loss: 0.010084139183163643
step: 370, loss: 0.0002606409543659538
step: 380, loss: 0.001886266632936895
step: 390, loss: 0.018549809232354164
step: 400, loss: 0.00019635033095255494
step: 410, loss: 0.004759388044476509
step: 420, loss: 0.0024585877545177937
step: 430, loss: 0.0025153413880616426
step: 440, loss: 0.0008433407638221979
step: 450, loss: 0.000759886868763715
step: 460, loss: 0.004024455789476633
step: 470, loss: 0.01588137447834015
step: 480, loss: 0.000956988544203341
step: 490, loss: 0.0019566507544368505
step: 500, loss: 0.0006405600579455495
step: 510, loss: 0.00048702064668759704
step: 520, loss: 4.850848563364707e-05
step: 530, loss: 0.0006307841977104545
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9223616922361693, f1=0.9178338001867413,
best_f1=0.9199063231850116
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006871194927953184

```

step: 10, loss: 0.02002161182463169  
step: 20, loss: 0.0014835660113021731  
step: 30, loss: 0.0002644349879119545  
step: 40, loss: 0.0017781922360882163  
step: 50, loss: 0.0003514629788696766  
step: 60, loss: 0.0380958653986454  
step: 70, loss: 0.00669792341068387  
step: 80, loss: 0.0018814422655850649  
step: 90, loss: 0.005440158303827047  
step: 100, loss: 0.004841462709009647  
step: 110, loss: 0.00445340434089303  
step: 120, loss: 0.000541717279702425  
step: 130, loss: 0.004342155531048775  
step: 140, loss: 0.0005335049354471266  
step: 150, loss: 0.005395867396146059  
step: 160, loss: 0.0009630785207264125  
step: 170, loss: 0.00034452942782081664  
step: 180, loss: 0.008866213262081146  
step: 190, loss: 0.0029645850881934166  
step: 200, loss: 0.0006433024536818266  
step: 210, loss: 0.0021340171806514263  
step: 220, loss: 0.00027813142514787614  
step: 230, loss: 0.002238463843241334  
step: 240, loss: 0.00013989585568197072  
step: 250, loss: 0.0004058342310599983  
step: 260, loss: 0.004883377347141504  
step: 270, loss: 4.908238770440221e-05  
step: 280, loss: 0.00019945956591982394  
step: 290, loss: 0.00019607858848758042  
step: 300, loss: 0.0009383413125760853  
step: 310, loss: 0.0002927801397163421  
step: 320, loss: 7.893388828961179e-05  
step: 330, loss: 0.0002072053321171552  
step: 340, loss: 0.008038484491407871  
step: 350, loss: 0.0008117732941173017  
step: 360, loss: 0.001199539634399116  
step: 370, loss: 0.00026026403065770864  
step: 380, loss: 0.010641712695360184  
step: 390, loss: 0.0006109203677624464  
step: 400, loss: 0.00017631248920224607  
step: 410, loss: 0.039203349500894547  
step: 420, loss: 0.00042499465052969754  
step: 430, loss: 0.00346207432448864  
step: 440, loss: 0.004306027665734291  
step: 450, loss: 0.0004498165799304843  
step: 460, loss: 0.0037993951700627804  
step: 470, loss: 0.001015337067656219  
step: 480, loss: 0.00019991259614471346



```

step: 490, loss: 0.0009220772190019488
step: 500, loss: 0.009333650581538677
step: 510, loss: 6.03534426772967e-05
step: 520, loss: 0.014692865312099457
step: 530, loss: 0.0012270845472812653
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9239332096474954, f1=0.9193697868396664,
best_f1=0.9199063231850116
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002397872303845361
step: 10, loss: 0.0003301848191767931
step: 20, loss: 0.025882385671138763
step: 30, loss: 0.12007461488246918
step: 40, loss: 9.413726365892217e-05
step: 50, loss: 0.0010194511851295829
step: 60, loss: 0.002080423291772604
step: 70, loss: 0.004983247723430395
step: 80, loss: 0.06928877532482147
step: 90, loss: 0.005162965971976519
step: 100, loss: 5.918569513596594e-05
step: 110, loss: 0.004060889594256878
step: 120, loss: 0.005008208565413952
step: 130, loss: 0.0012107546208426356
step: 140, loss: 0.00013075537572149187
step: 150, loss: 0.00022718298714607954
step: 160, loss: 5.0034439482260495e-05
step: 170, loss: 0.0007827767403796315
step: 180, loss: 0.0001871745625976473
step: 190, loss: 0.00013402877084445208
step: 200, loss: 0.002890690928325057
step: 210, loss: 0.0016091946745291352
step: 220, loss: 0.0001584464916959405
step: 230, loss: 0.009197494946420193
step: 240, loss: 0.000174427215824835
step: 250, loss: 0.0440317764878273
step: 260, loss: 0.000198870780877769
step: 270, loss: 0.008860054425895214
step: 280, loss: 0.00017307056987192482
step: 290, loss: 0.00013629575551021844
step: 300, loss: 0.00047018329496495426
step: 310, loss: 9.777351806405932e-05
step: 320, loss: 0.002184745389968157
step: 330, loss: 0.0015987324295565486
step: 340, loss: 5.329529449227266e-05

```

```

step: 350, loss: 0.00967919360846281
step: 360, loss: 9.425233292859048e-05
step: 370, loss: 0.0013738877605646849
step: 380, loss: 0.0013243985595181584
step: 390, loss: 0.0005694939754903316
step: 400, loss: 0.0003797227400355041
step: 410, loss: 0.00019271635392215103
step: 420, loss: 0.00040800016722641885
step: 430, loss: 0.0003795150842051953
step: 440, loss: 0.006133448798209429
step: 450, loss: 0.0004410461406223476
step: 460, loss: 5.97029211348854e-05
step: 470, loss: 0.009214241988956928
step: 480, loss: 0.0016351983649656177
step: 490, loss: 0.0003459904692135751
step: 500, loss: 0.002967266133055091
step: 510, loss: 0.00029332475969567895
step: 520, loss: 0.0001542185782454908
step: 530, loss: 0.00047372811241075397
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 13: dev_f1=0.9264909847434121, f1=0.9205909510618652,
best_f1=0.9205909510618652
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.760972559684888e-05
step: 10, loss: 6.71559973852709e-05
step: 20, loss: 0.00031906182994134724
step: 30, loss: 0.012878906913101673
step: 40, loss: 0.0005009949672967196
step: 50, loss: 0.005331366788595915
step: 60, loss: 0.00013393790868576616
step: 70, loss: 0.08837555348873138
step: 80, loss: 6.305122224148363e-05
step: 90, loss: 0.00017559004481881857
step: 100, loss: 0.029857417568564415
step: 110, loss: 8.585231262259185e-05
step: 120, loss: 0.001453914912417531
step: 130, loss: 0.0006545270443893969
step: 140, loss: 0.0013540334766730666
step: 150, loss: 0.0042122467420995235
step: 160, loss: 0.00071982346707955

```

```

step: 170, loss: 0.002523301634937525
step: 180, loss: 0.002706753322854638
step: 190, loss: 0.0008527816389687359
step: 200, loss: 0.00014479584933724254
step: 210, loss: 0.0006351307965815067
step: 220, loss: 0.002866321476176381
step: 230, loss: 0.0019699293188750744
step: 240, loss: 0.004838378168642521
step: 250, loss: 0.004822825081646442
step: 260, loss: 3.369788100826554e-05
step: 270, loss: 0.0010373778641223907
step: 280, loss: 0.0004060637147631496
step: 290, loss: 0.0011063497513532639
step: 300, loss: 0.000260234868619591
step: 310, loss: 0.00046808322076685727
step: 320, loss: 0.0003066129283979535
step: 330, loss: 0.0003675088519230485
step: 340, loss: 0.00017562629363965243
step: 350, loss: 0.0033748215064406395
step: 360, loss: 0.039348166435956955
step: 370, loss: 0.0009820484556257725
step: 380, loss: 8.115866512525827e-05
step: 390, loss: 0.0016056143213063478
step: 400, loss: 0.00022807286586612463
step: 410, loss: 0.001226454391144216
step: 420, loss: 0.0025197756476700306
step: 430, loss: 3.087802178924903e-05
step: 440, loss: 3.306050712126307e-05
step: 450, loss: 4.489218918024562e-05
step: 460, loss: 0.00017947060405276716
step: 470, loss: 0.00047480958164669573
step: 480, loss: 5.4068495956016704e-05
step: 490, loss: 0.0007742180023342371
step: 500, loss: 0.18665899336338043
step: 510, loss: 0.0028163515962660313
step: 520, loss: 7.077163900248706e-05
step: 530, loss: 0.0019966710824519396
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9241379310344828, f1=0.9232183908045977,
best_f1=0.9205909510618652
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00051738420734182
step: 10, loss: 0.002840134548023343
step: 20, loss: 0.000533728045411408

```

step: 30, loss: 0.00011298718163743615  
step: 40, loss: 0.00019731299835257232  
step: 50, loss: 0.0001642649876885116  
step: 60, loss: 8.114507363643497e-05  
step: 70, loss: 0.0002178447030019015  
step: 80, loss: 9.200126805808395e-05  
step: 90, loss: 2.108859371219296e-05  
step: 100, loss: 0.0007997548673301935  
step: 110, loss: 0.013818648643791676  
step: 120, loss: 0.005523744039237499  
step: 130, loss: 0.0001837956515373662  
step: 140, loss: 8.284962677862495e-05  
step: 150, loss: 0.000158819486387074  
step: 160, loss: 0.001169699477031827  
step: 170, loss: 0.0002677692682482302  
step: 180, loss: 0.0044761826284229755  
step: 190, loss: 0.0005615871050395072  
step: 200, loss: 0.0006266958080232143  
step: 210, loss: 0.0034103537909686565  
step: 220, loss: 0.0005622111493721604  
step: 230, loss: 0.00233081285841763  
step: 240, loss: 0.0003698783984873444  
step: 250, loss: 0.0011778655461966991  
step: 260, loss: 0.00028271734481677413  
step: 270, loss: 0.05341098830103874  
step: 280, loss: 0.00010463916987646371  
step: 290, loss: 0.0014269300736486912  
step: 300, loss: 0.0013404544442892075  
step: 310, loss: 0.0007389613310806453  
step: 320, loss: 0.00403369078412652  
step: 330, loss: 0.1559978723526001  
step: 340, loss: 0.0008113491930998862  
step: 350, loss: 0.0006619102787226439  
step: 360, loss: 0.0002161260781576857  
step: 370, loss: 0.0034492043778300285  
step: 380, loss: 0.00039840006502345204  
step: 390, loss: 0.0005018698866479099  
step: 400, loss: 9.711947495816275e-05  
step: 410, loss: 0.0004266907926648855  
step: 420, loss: 0.0006130092078819871  
step: 430, loss: 5.6589684390928596e-05  
step: 440, loss: 0.005632100161164999  
step: 450, loss: 0.003438961924985051  
step: 460, loss: 0.0010147084249183536  
step: 470, loss: 0.00016072046128101647  
step: 480, loss: 0.0010521805379539728  
step: 490, loss: 0.00010318720887880772  
step: 500, loss: 5.1366674597375095e-05

```

step: 510, loss: 0.00011459488450782374
step: 520, loss: 0.00013347713684197515
step: 530, loss: 0.0002131390356225893
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9240681086056144, f1=0.9232889297197979,
best_f1=0.9205909510618652
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:14, 392.58it/s]
load_f1 = 0.922089825847846
real_f1 = 0.9207512597343105
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:11, 389.83it/s]

```

## 2 BASELINE

### 2.1 BASELINE STRUCTURED

#### 2.1.1 Beer - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/Beer \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm distilbert \
  --fp16 \
  --save_model

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Beer.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/Beer \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm distilbert \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/
```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```
enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic
```

Processing user overrides (additional kwargs that are not None)...

```

After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.8721616268157959
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.30000000000000004, f1=0.23076923076923075,
best_f1=0.23076923076923075
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3534112274646759
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.3414634146341463, f1=0.2758620689655172,
best_f1=0.2758620689655172

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3546009361743927
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.3333333333333337, f1=0.3225806451612903,
best_f1=0.2758620689655172
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.33347412943840027
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.3157894736842105, f1=0.2790697674418604,
best_f1=0.2758620689655172
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22852222621440887
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.3835616438356164, f1=0.25641025641025644,
best_f1=0.25641025641025644
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2621252238750458
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 6: dev_f1=0.4126984126984127, f1=0.2857142857142857,
best_f1=0.2857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1848420649766922

```



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.44067796610169485, f1=0.30303030303030304,
best_f1=0.30303030303030304
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.34902989864349365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.44067796610169485, f1=0.35714285714285715,
best_f1=0.30303030303030304
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16633544862270355
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.41791044776119407, f1=0.38095238095238093,
best_f1=0.30303030303030304
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2088109701871872
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 10: dev_f1=0.5142857142857143, f1=0.41666666666666663,
best_f1=0.41666666666666663
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26125386357307434
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.4864864864864865, f1=0.4, best_f1=0.41666666666666663

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12905998528003693
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 12: dev_f1=0.5405405405405405, f1=0.41666666666666663,
best_f1=0.41666666666666663
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11019492149353027
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 13: dev_f1=0.5714285714285714, f1=0.41666666666666663,
best_f1=0.41666666666666663
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1521023064851761
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.5555555555555556, f1=0.41666666666666663,
best_f1=0.41666666666666663
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1909942775964737
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.5555555555555556, f1=0.41666666666666663,
best_f1=0.41666666666666663
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

```

```

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
91it [00:00, 142951.93it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
load_f1 = 0.43750000000000006
real_f1 = 0.5263157894736842
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:11, 397.53it/s]

```

### 2.1.2 DBLP-ACM - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
--task Structured/DBLP-ACM \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \

```

```

--finetuning \
--lm distilbert \
--fp16 \
--save_model

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-ACM.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/DBLP-ACM \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm distilbert \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0  
/usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`.

In PyTorch 1.1.0 and later, you should call them in the opposite order: ``optimizer.step()`` before ``lr_scheduler.step()``. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

```
"https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
```

```
step: 0, loss: 0.8085125088691711
```

```
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
```

```
step: 10, loss: 0.4748813509941101
```

```
step: 20, loss: 0.5985168218612671
```

```
step: 30, loss: 0.4722555875778198
```

```
step: 40, loss: 0.2896931767463684
```

```
step: 50, loss: 0.25400590896606445
```

```
step: 60, loss: 0.10737159848213196
```

```
step: 70, loss: 0.22509099543094635
```

```
step: 80, loss: 0.1680591106414795
```

```
step: 90, loss: 0.008161083795130253
```

```
step: 100, loss: 0.033381540328264236
```

```
step: 110, loss: 0.07438530027866364
```

```
step: 120, loss: 0.032682035118341446
```

```
step: 130, loss: 0.10853473842144012
```

```
step: 140, loss: 0.026370344683527946
```

```
step: 150, loss: 0.016919277608394623
```

```
step: 160, loss: 0.21524231135845184
```

```
step: 170, loss: 0.0036334244068711996
```

```
step: 180, loss: 0.007871991954743862
```

```
step: 190, loss: 0.03851327672600746
```

```
step: 200, loss: 0.04099653288722038
```

```
step: 210, loss: 0.010943367145955563
```

```
step: 220, loss: 0.006849810481071472
```

```
step: 230, loss: 0.008278719149529934
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
```

```
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

```
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
```

```
warnings.warn(SAVE_STATE_WARNING, UserWarning)
```

```
epoch 1: dev_f1=0.9765363128491621, f1=0.9774774774774775,
```

```
best_f1=0.9774774774774775
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.02047174796462059
```

```
step: 10, loss: 0.0010321809677407146
```

```
step: 20, loss: 0.004025481175631285
```

```
step: 30, loss: 0.0017344314837828279
```

```

step: 40, loss: 0.0044727567583322525
step: 50, loss: 0.00220456812530756
step: 60, loss: 0.005451207980513573
step: 70, loss: 0.008012466132640839
step: 80, loss: 0.0029979576356709003
step: 90, loss: 0.015862297266721725
step: 100, loss: 0.017546474933624268
step: 110, loss: 0.004959168843924999
step: 120, loss: 0.012048514559864998
step: 130, loss: 0.003120872424915433
step: 140, loss: 0.012735013850033283
step: 150, loss: 0.006008841563016176
step: 160, loss: 0.11424148082733154
step: 170, loss: 0.10657460987567902
step: 180, loss: 0.0020744360517710447
step: 190, loss: 0.17999979853630066
step: 200, loss: 0.01011161133646965
step: 210, loss: 0.017028555274009705
step: 220, loss: 0.0026781216729432344
step: 230, loss: 0.06755878031253815
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9797752808988766, f1=0.9707207207207207,
best_f1=0.9707207207207207
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016475651413202286
step: 10, loss: 0.06770236790180206
step: 20, loss: 0.006291077006608248
step: 30, loss: 0.07605340331792831
step: 40, loss: 0.03806690499186516
step: 50, loss: 0.021136246621608734
step: 60, loss: 0.0008211340173147619
step: 70, loss: 0.0021375229116529226
step: 80, loss: 0.07151398062705994
step: 90, loss: 0.007902721874415874
step: 100, loss: 0.002445526421070099
step: 110, loss: 0.029628347605466843
step: 120, loss: 0.014939668588340282
step: 130, loss: 0.0017144359881058335
step: 140, loss: 0.0006363497232086957
step: 150, loss: 0.0012181611964479089

```

```

step: 160, loss: 0.0013559620128944516
step: 170, loss: 0.010521423071622849
step: 180, loss: 0.0061499737203121185
step: 190, loss: 0.0021059163846075535
step: 200, loss: 0.04135611280798912
step: 210, loss: 0.00257158768363297
step: 220, loss: 0.004723150748759508
step: 230, loss: 0.19249074161052704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.9787234042553192, f1=0.9819413092550789,
best_f1=0.9707207207207207
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0595967099070549
step: 10, loss: 0.0016595941269770265
step: 20, loss: 0.00267406995408237
step: 30, loss: 0.0016810783417895436
step: 40, loss: 0.006449727341532707
step: 50, loss: 0.004050272051244974
step: 60, loss: 0.001520955003798008
step: 70, loss: 0.005955889821052551
step: 80, loss: 0.06761228293180466
step: 90, loss: 0.01014786772429943
step: 100, loss: 0.022806422784924507
step: 110, loss: 0.004553150851279497
step: 120, loss: 0.025377556681632996
step: 130, loss: 0.008919025771319866
step: 140, loss: 0.012065205723047256
step: 150, loss: 0.058028437197208405
step: 160, loss: 0.0009009675704874098
step: 170, loss: 0.017089229077100754
step: 180, loss: 0.03033575415611267
step: 190, loss: 0.015368820168077946
step: 200, loss: 0.004068406764417887
step: 210, loss: 0.0009897849522531033
step: 220, loss: 0.0021608888637274504
step: 230, loss: 0.0006374374497681856
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.984304932735426, f1=0.9798206278026906,

```

```

best_f1=0.9798206278026906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004017926403321326
step: 10, loss: 0.0003599741030484438
step: 20, loss: 0.0006007991032674909
step: 30, loss: 0.00018858944531530142
step: 40, loss: 0.0014392873272299767
step: 50, loss: 0.0007891854620538652
step: 60, loss: 0.0004733466194011271
step: 70, loss: 0.0036484987940639257
step: 80, loss: 0.0006864310707896948
step: 90, loss: 0.001382289920002222
step: 100, loss: 0.0009756509680300951
step: 110, loss: 0.00043949426617473364
step: 120, loss: 0.049932513386011124
step: 130, loss: 0.0008042609551921487
step: 140, loss: 0.0003438251151237637
step: 150, loss: 0.00019227404845878482
step: 160, loss: 0.0032355310395359993
step: 170, loss: 0.04821375012397766
step: 180, loss: 0.001017497619614005
step: 190, loss: 0.0009973073611035943
step: 200, loss: 0.007757660932838917
step: 210, loss: 0.0012589647667482495
step: 220, loss: 0.0011320413323119283
step: 230, loss: 0.0002496821980457753
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.9853438556933484, f1=0.9797752808988766,
best_f1=0.9797752808988766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00048306898679584265
step: 10, loss: 0.0021520834416151047
step: 20, loss: 0.00045890980982221663
step: 30, loss: 0.00021203650976531208
step: 40, loss: 0.0005410539451986551
step: 50, loss: 0.0010319034336134791
step: 60, loss: 0.006619269959628582
step: 70, loss: 0.001005207421258092

```



```

step: 80, loss: 0.0002200750313932076
step: 90, loss: 0.0001925759861478582
step: 100, loss: 0.00019910780247300863
step: 110, loss: 0.006633858662098646
step: 120, loss: 0.000234541657846421
step: 130, loss: 0.0021941731683909893
step: 140, loss: 0.0027718190103769302
step: 150, loss: 0.0008723746286705136
step: 160, loss: 0.037161316722631454
step: 170, loss: 0.0001440008491044864
step: 180, loss: 0.0009043082245625556
step: 190, loss: 0.0016938479384407401
step: 200, loss: 0.001151863718405366
step: 210, loss: 0.0011309247929602861
step: 220, loss: 0.001279782853089273
step: 230, loss: 0.000806552532594651
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9830890642615557, f1=0.9808342728297633,
best_f1=0.9797752808988766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020502416417002678
step: 10, loss: 0.0002757779147941619
step: 20, loss: 0.0013289402704685926
step: 30, loss: 0.00031327037140727043
step: 40, loss: 0.004248670767992735
step: 50, loss: 0.00012418015103321522
step: 60, loss: 0.00035719270817935467
step: 70, loss: 0.0006236853660084307
step: 80, loss: 0.0003034131077583879
step: 90, loss: 0.00042929203482344747
step: 100, loss: 0.0016265427693724632
step: 110, loss: 0.0065849763341248035
step: 120, loss: 0.00207247375510633
step: 130, loss: 0.02311014197766781
step: 140, loss: 0.0003845977771561593
step: 150, loss: 0.0017918009543791413
step: 160, loss: 0.0005141549045220017
step: 170, loss: 0.00024372337793465704
step: 180, loss: 0.0003387940814718604
step: 190, loss: 0.0015972849214449525
step: 200, loss: 0.010420096106827259
step: 210, loss: 0.0005286165396682918
step: 220, loss: 0.0009737377404235303
step: 230, loss: 0.025375129655003548

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.976324689966178, f1=0.9774266365688488,
best_f1=0.9797752808988766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02890835888683796
step: 10, loss: 0.02465656027197838
step: 20, loss: 0.0031144001986831427
step: 30, loss: 0.003922830335795879
step: 40, loss: 0.0006159406621009111
step: 50, loss: 0.0007737056002952158
step: 60, loss: 0.0003658013592939824
step: 70, loss: 0.0004259505949448794
step: 80, loss: 0.001177517930045724
step: 90, loss: 0.00029916560743004084
step: 100, loss: 0.000768858939409256
step: 110, loss: 0.00069397286279127
step: 120, loss: 0.004654130898416042
step: 130, loss: 0.004894668236374855
step: 140, loss: 0.004082700237631798
step: 150, loss: 0.042305123060941696
step: 160, loss: 0.0006894321995787323
step: 170, loss: 0.0011670540552586317
step: 180, loss: 0.0034791689831763506
step: 190, loss: 0.0010385829955339432
step: 200, loss: 0.0010968855349346995
step: 210, loss: 0.00036264871596358716
step: 220, loss: 0.004970100708305836
step: 230, loss: 0.008923647925257683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9808342728297633, f1=0.9739524348810873,
best_f1=0.9797752808988766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02473696507513523
step: 10, loss: 0.00029787016683258116
step: 20, loss: 0.007894618436694145
step: 30, loss: 0.00018914883548859507
step: 40, loss: 0.00010410523100290447
step: 50, loss: 0.00010516872862353921
step: 60, loss: 0.0001761903549777344
step: 70, loss: 0.0004160947573836893

```

```

step: 80, loss: 0.012066812254488468
step: 90, loss: 0.00047791891847737134
step: 100, loss: 0.0003332322812639177
step: 110, loss: 0.001972471596673131
step: 120, loss: 0.03218965604901314
step: 130, loss: 0.0006013744859956205
step: 140, loss: 0.00025190520682372153
step: 150, loss: 0.00013327656779438257
step: 160, loss: 0.0017330803675577044
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 170, loss: 0.00020782741194125265
step: 180, loss: 0.0003355265944264829
step: 190, loss: 0.00038533948827534914
step: 200, loss: 0.0004537898057606071
step: 210, loss: 0.0003464054607320577
step: 220, loss: 0.02640906535089016
step: 230, loss: 0.010105096735060215
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9842342342342343, f1=0.9807474518686297,
best_f1=0.9797752808988766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00035152919008396566
step: 10, loss: 0.0004288179916329682
step: 20, loss: 0.0004182924167253077
step: 30, loss: 0.00026703529874794185
step: 40, loss: 0.00011745913798222318
step: 50, loss: 0.0005457851802930236
step: 60, loss: 0.022804107517004013
step: 70, loss: 0.00018760759849101305
step: 80, loss: 0.00015567467198707163
step: 90, loss: 0.00024236664467025548
step: 100, loss: 0.00023179473646450788
step: 110, loss: 0.0002690241381060332
step: 120, loss: 0.012246071361005306
step: 130, loss: 0.0001433858269592747
step: 140, loss: 0.02845720574259758
step: 150, loss: 0.00020812448929063976
step: 160, loss: 0.000514196464791894
step: 170, loss: 0.0009008926572278142
step: 180, loss: 0.05188075453042984
step: 190, loss: 0.0008548050536774099
step: 200, loss: 0.000193811472854577
step: 210, loss: 0.0006399106932803988
step: 220, loss: 0.0004623596032615751

```

```

step: 230, loss: 0.00024468981428071856
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9841986455981941, f1=0.983050847457627,
best_f1=0.9797752808988766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00033497647382318974
step: 10, loss: 0.0005503032589331269
step: 20, loss: 0.00013983876851852983
step: 30, loss: 0.0003269874432589859
step: 40, loss: 0.0010492542060092092
step: 50, loss: 0.0005011037574149668
step: 60, loss: 0.0034890719689428806
step: 70, loss: 0.0006023267051205039
step: 80, loss: 0.0005388554418459535
step: 90, loss: 0.00028768289485014975
step: 100, loss: 0.0003988573735114187
step: 110, loss: 0.00023480683739762753
step: 120, loss: 0.00029433888266794384
step: 130, loss: 0.00015310445451177657
step: 140, loss: 0.00037381594302132726
step: 150, loss: 0.00015227028052322567
step: 160, loss: 0.00018139580788556486
step: 170, loss: 0.01010061800479889
step: 180, loss: 0.0003325300058349967
step: 190, loss: 0.0008466975996270776
step: 200, loss: 0.0002904606517404318
step: 210, loss: 0.0002360329672228545
step: 220, loss: 0.0002081843267660588
step: 230, loss: 0.00027722519007511437
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9853438556933484, f1=0.9841986455981941,
best_f1=0.9797752808988766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001024798839353025
step: 10, loss: 0.00033189920941367745
step: 20, loss: 0.00012441622675396502
step: 30, loss: 0.00043906611972488463
step: 40, loss: 0.00011653727415250614
step: 50, loss: 0.00016875698929652572
step: 60, loss: 0.00020668760407716036

```

```

step: 70, loss: 0.0001423798967152834
step: 80, loss: 8.675410936120898e-05
step: 90, loss: 0.00013432270498014987
step: 100, loss: 0.00011312448623357341
step: 110, loss: 0.0002005299465963617
step: 120, loss: 0.00016148528084158897
step: 130, loss: 0.000285852060187608
step: 140, loss: 0.00012892046652268618
step: 150, loss: 0.00010775317787192762
step: 160, loss: 0.0362466536462307
step: 170, loss: 0.0003038886352442205
step: 180, loss: 0.00014231052773538977
step: 190, loss: 0.00020646881603170186
step: 200, loss: 0.000750677427276969
step: 210, loss: 0.00012282220995984972
step: 220, loss: 0.0002745011879596859
step: 230, loss: 0.0001370298268739134
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9841986455981941, f1=0.9841986455981941,
best_f1=0.9797752808988766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017281736654695123
step: 10, loss: 0.0010352024110034108
step: 20, loss: 0.0002584734174888581
step: 30, loss: 0.028231576085090637
step: 40, loss: 0.00021015464153606445
step: 50, loss: 0.00022674573119729757
step: 60, loss: 0.0004902068176306784
step: 70, loss: 0.0003692824684549123
step: 80, loss: 8.13961451058276e-05
step: 90, loss: 8.228594379033893e-05
step: 100, loss: 0.0001978971267817542
step: 110, loss: 0.0001892777654575184
step: 120, loss: 0.00017484890122432262
step: 130, loss: 0.00012263740063644946
step: 140, loss: 0.00019003338820766658
step: 150, loss: 0.03704475611448288
step: 160, loss: 0.0001557734503876418
step: 170, loss: 0.00013811845565214753
step: 180, loss: 0.0001598442322574556
step: 190, loss: 6.694121839245781e-05
step: 200, loss: 0.00017100386321544647
step: 210, loss: 0.00011586866457946599
step: 220, loss: 6.655402830801904e-05

```

```

step: 230, loss: 8.370757132070139e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9831649831649831, f1=0.9830890642615557,
best_f1=0.9797752808988766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011418773647164926
step: 10, loss: 5.172467353986576e-05
step: 20, loss: 0.019206088036298752
step: 30, loss: 0.00016093773592729121
step: 40, loss: 7.658881804673001e-05
step: 50, loss: 8.739499025978148e-05
step: 60, loss: 5.508307003765367e-05
step: 70, loss: 0.00011272540723439306
step: 80, loss: 0.0003290930762887001
step: 90, loss: 0.00014320859918370843
step: 100, loss: 0.005819191224873066
step: 110, loss: 0.00012711896852124482
step: 120, loss: 0.004781566094607115
step: 130, loss: 0.0002282575296703726
step: 140, loss: 0.00012162497296230868
step: 150, loss: 0.0002526272728573531
step: 160, loss: 0.0001213427894981578
step: 170, loss: 0.0001024417724693194
step: 180, loss: 0.0004287199699319899
step: 190, loss: 0.0031638762447983027
step: 200, loss: 7.353644468821585e-05
step: 210, loss: 0.0007294326205737889
step: 220, loss: 0.00016918595065362751
step: 230, loss: 6.731232133461162e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9830890642615557, f1=0.9830890642615557,
best_f1=0.9797752808988766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011025171988876536
step: 10, loss: 0.00011499373795231804
step: 20, loss: 0.00011044296843465418
step: 30, loss: 0.00043375161476433277
step: 40, loss: 0.00024888344341889024
step: 50, loss: 0.00030901882564648986
step: 60, loss: 8.478320523863658e-05

```

```

step: 70, loss: 0.015365583822131157
step: 80, loss: 0.0001156675789388828
step: 90, loss: 5.11252655996941e-05
step: 100, loss: 0.0001549895096104592
step: 110, loss: 9.778504318092018e-05
step: 120, loss: 0.00013849746028427035
step: 130, loss: 0.00012175490701338276
step: 140, loss: 0.0026859454810619354
step: 150, loss: 0.00014433501928579062
step: 160, loss: 0.00012484873877838254
step: 170, loss: 6.589016993530095e-05
step: 180, loss: 0.00011448265286162496
step: 190, loss: 0.00515540037304163
step: 200, loss: 0.0001184672219096683
step: 210, loss: 0.0334509052336216
step: 220, loss: 0.00011263372289249673
step: 230, loss: 0.00015599728794768453
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9842696629213483, f1=0.9830890642615557,
best_f1=0.9797752808988766
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
2473it [00:07, 326.11it/s]
load_f1 = 0.9842342342342343
real_f1 = 0.9853438556933484
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
4398it [00:11, 396.18it/s]
```

### 2.1.3 DBLP-GoogleScholar - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
--task Structured/DBLP-GoogleScholar \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm distilbert \
--fp16 \
--save_model

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-GoogleScholar.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
--task Structured/DBLP-GoogleScholar \
--input_path input/input_small.jsonl \
--output_path output/output_small.jsonl \
--lm distilbert \
--use_gpu \
--fp16 \
--checkpoint_path checkpoints/
```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.



Defaults for this optimization level are:

```
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback.

Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0  
/usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`.

In PyTorch 1.1.0 and later, you should call them in the opposite order:

`optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

"<https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>",  
UserWarning)

```
step: 0, loss: 0.8002065420150757
step: 10, loss: 0.4267975687980652
step: 20, loss: 0.49769723415374756
step: 30, loss: 0.4439997673034668
step: 40, loss: 0.35463759303092957
step: 50, loss: 0.26660966873168945
step: 60, loss: 0.21807336807250977
step: 70, loss: 0.13226740062236786
step: 80, loss: 0.21895605325698853
step: 90, loss: 0.16788698732852936
step: 100, loss: 0.3515928387641907
step: 110, loss: 0.07736985385417938
step: 120, loss: 0.13617566227912903
step: 130, loss: 0.05361523479223251
step: 140, loss: 0.3043229281902313
```

```

step: 150, loss: 0.02669634483754635
step: 160, loss: 0.0943358913064003
step: 170, loss: 0.23846641182899475
step: 180, loss: 0.14882877469062805
step: 190, loss: 0.05463230609893799
step: 200, loss: 0.11359523981809616
step: 210, loss: 0.1852668821811676
step: 220, loss: 0.09237606823444366
step: 230, loss: 0.11623446643352509
step: 240, loss: 0.08861218392848969
step: 250, loss: 0.08632533252239227
step: 260, loss: 0.0632948949933052
step: 270, loss: 0.028150277212262154
step: 280, loss: 0.20267130434513092
step: 290, loss: 0.112989641726017
step: 300, loss: 0.12416447699069977
step: 310, loss: 0.19195419549942017
step: 320, loss: 0.09965866804122925
step: 330, loss: 0.08796413242816925
step: 340, loss: 0.14412665367126465
step: 350, loss: 0.026450572535395622
step: 360, loss: 0.12780986726284027
step: 370, loss: 0.18666972219944
step: 380, loss: 0.2357083410024643
step: 390, loss: 0.08093221485614777
step: 400, loss: 0.05263017490506172
step: 410, loss: 0.0519917793571949
step: 420, loss: 0.08460874110460281
step: 430, loss: 0.0815095603466034
step: 440, loss: 0.10165902972221375
step: 450, loss: 0.09981927275657654
step: 460, loss: 0.05324644595384598
step: 470, loss: 0.305133193731308
step: 480, loss: 0.18143904209136963
step: 490, loss: 0.04644317924976349
step: 500, loss: 0.030090274289250374
step: 510, loss: 0.08524325489997864
step: 520, loss: 0.105848528444767
step: 530, loss: 0.057645030319690704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9180778032036614, f1=0.9157942649066909,
best_f1=0.9157942649066909

```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.
```

```
    warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.15160012245178223  
step: 10, loss: 0.15832816064357758  
step: 20, loss: 0.16595254838466644  
step: 30, loss: 0.05024387314915657  
step: 40, loss: 0.003640955314040184  
step: 50, loss: 0.17467404901981354  
step: 60, loss: 0.09671672433614731  
step: 70, loss: 0.13312216103076935  
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0  
step: 80, loss: 0.06383761763572693  
step: 90, loss: 0.050290606915950775  
step: 100, loss: 0.33670753240585327  
step: 110, loss: 0.0606452152132988  
step: 120, loss: 0.055308498442173004  
step: 130, loss: 0.05012299120426178  
step: 140, loss: 0.03269268944859505  
step: 150, loss: 0.01698455587029457  
step: 160, loss: 0.26555687189102173  
step: 170, loss: 0.09472818672657013  
step: 180, loss: 0.004351639188826084  
step: 190, loss: 0.03071717359125614  
step: 200, loss: 0.0562899112701416  
step: 210, loss: 0.0270057525485754  
step: 220, loss: 0.12304212152957916  
step: 230, loss: 0.03733126074075699  
step: 240, loss: 0.09896381199359894  
step: 250, loss: 0.052609313279390335  
step: 260, loss: 0.006010915618389845  
step: 270, loss: 0.045545682311058044  
step: 280, loss: 0.2797415852546692  
step: 290, loss: 0.06025014817714691  
step: 300, loss: 0.020525336265563965  
step: 310, loss: 0.0658608004450798  
step: 320, loss: 0.050761785358190536  
step: 330, loss: 0.03706003352999687  
step: 340, loss: 0.015887174755334854  
step: 350, loss: 0.058640800416469574  
step: 360, loss: 0.14208421111106873  
step: 370, loss: 0.009577385149896145  
step: 380, loss: 0.059344954788684845  
step: 390, loss: 0.05128313601016998  
step: 400, loss: 0.043982457369565964  
step: 410, loss: 0.004592980723828077  
step: 420, loss: 0.08157429844141006  
step: 430, loss: 0.029476836323738098
```

```

step: 440, loss: 0.012955611571669579
step: 450, loss: 0.0059177144430577755
step: 460, loss: 0.2176322489976883
step: 470, loss: 0.033255524933338165
step: 480, loss: 0.21374468505382538
step: 490, loss: 0.039219170808792114
step: 500, loss: 0.045475345104932785
step: 510, loss: 0.0460202619433403
step: 520, loss: 0.023439444601535797
step: 530, loss: 0.11094816774129868
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.923286427598729, f1=0.9198350893266148,
best_f1=0.9198350893266148
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06674977391958237
step: 10, loss: 0.05396347865462303
step: 20, loss: 0.05968949571251869
step: 30, loss: 0.15303446352481842
step: 40, loss: 0.002745706122368574
step: 50, loss: 0.007824547588825226
step: 60, loss: 0.00427551195025444
step: 70, loss: 0.08319824188947678
step: 80, loss: 0.0027917989064007998
step: 90, loss: 0.15099188685417175
step: 100, loss: 0.03174692019820213
step: 110, loss: 0.08136164397001266
step: 120, loss: 0.01271207444369793
step: 130, loss: 0.027564838528633118
step: 140, loss: 0.029981205239892006
step: 150, loss: 0.03122718445956707
step: 160, loss: 0.0023645071778446436
step: 170, loss: 0.005480986554175615
step: 180, loss: 0.0039007975719869137
step: 190, loss: 0.0053145429119467735
step: 200, loss: 0.013293074443936348
step: 210, loss: 0.0529015026986599
step: 220, loss: 0.024890122935175896
step: 230, loss: 0.07360867410898209
step: 240, loss: 0.007087585050612688
step: 250, loss: 0.03321558237075806

```

```

step: 260, loss: 0.003017859300598502
step: 270, loss: 0.0022267058957368135
step: 280, loss: 0.10587755590677261
step: 290, loss: 0.0477045476436615
step: 300, loss: 0.11386922001838684
step: 310, loss: 0.07982122898101807
step: 320, loss: 0.07687671482563019
step: 330, loss: 0.010814689099788666
step: 340, loss: 0.007311280816793442
step: 350, loss: 0.014362026005983353
step: 360, loss: 0.029402801766991615
step: 370, loss: 0.006094551645219326
step: 380, loss: 0.06428588926792145
step: 390, loss: 0.04374215379357338
step: 400, loss: 0.01776133105158806
step: 410, loss: 0.012670306488871574
step: 420, loss: 0.06064581498503685
step: 430, loss: 0.012210429646074772
step: 440, loss: 0.021895645186305046
step: 450, loss: 0.137827068567276
step: 460, loss: 0.0585903637111187
step: 470, loss: 0.006919478997588158
step: 480, loss: 0.009041291661560535
step: 490, loss: 0.016623716801404953
step: 500, loss: 0.09827812761068344
step: 510, loss: 0.05226903781294823
step: 520, loss: 0.019066788256168365
step: 530, loss: 0.006525238510221243
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.9288664525011474, f1=0.9296803652968036,
best_f1=0.9296803652968036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01033820677548647
step: 10, loss: 0.004821352660655975
step: 20, loss: 0.09457003325223923
step: 30, loss: 0.004834079183638096
step: 40, loss: 0.007256834767758846
step: 50, loss: 0.02979210764169693
step: 60, loss: 0.0003708404547069222
step: 70, loss: 0.007851776666939259

```

```
step: 80, loss: 0.0016936809988692403
step: 90, loss: 0.027444664388895035
step: 100, loss: 0.009916492737829685
step: 110, loss: 0.004954788368195295
step: 120, loss: 0.002611313248053193
step: 130, loss: 0.0050224727019667625
step: 140, loss: 0.0029883217066526413
step: 150, loss: 0.022198637947440147
step: 160, loss: 0.004352407529950142
step: 170, loss: 0.01303095556795597
step: 180, loss: 0.04293610900640488
step: 190, loss: 0.011491290293633938
step: 200, loss: 0.006225530989468098
step: 210, loss: 0.007995463907718658
step: 220, loss: 0.020076967775821686
step: 230, loss: 0.16066935658454895
step: 240, loss: 0.003961562179028988
step: 250, loss: 0.007534167729318142
step: 260, loss: 0.11921890079975128
step: 270, loss: 0.10358106344938278
step: 280, loss: 0.0779653862118721
step: 290, loss: 0.042313165962696075
step: 300, loss: 0.01236886065453291
step: 310, loss: 0.00480993278324604
step: 320, loss: 0.01702999696135521
step: 330, loss: 0.05998494476079941
step: 340, loss: 0.0034464066848158836
step: 350, loss: 0.001059086644090712
step: 360, loss: 0.0013506526593118906
step: 370, loss: 0.00903878640383482
step: 380, loss: 0.0023358638864010572
step: 390, loss: 0.04672225937247276
step: 400, loss: 0.011975506320595741
step: 410, loss: 0.006950058043003082
step: 420, loss: 0.014520843513309956
step: 430, loss: 0.0037156613543629646
step: 440, loss: 0.08360371738672256
step: 450, loss: 0.0449974350631237
step: 460, loss: 0.014567352831363678
step: 470, loss: 0.0022991413716226816
step: 480, loss: 0.001712334924377501
step: 490, loss: 0.08409860730171204
step: 500, loss: 0.029106101021170616
step: 510, loss: 0.03231707960367203
step: 520, loss: 0.009001331403851509
step: 530, loss: 0.003902824828401208
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```

warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.933579335793358, f1=0.9288354898336414,
best_f1=0.9288354898336414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015403113793581724
step: 10, loss: 0.0017754622967913747
step: 20, loss: 0.0005994127714075148
step: 30, loss: 0.03959392383694649
step: 40, loss: 0.026831287890672684
step: 50, loss: 0.04383594170212746
step: 60, loss: 0.0010477065807208419
step: 70, loss: 0.004148569423705339
step: 80, loss: 0.00031916238367557526
step: 90, loss: 0.03804351016879082
step: 100, loss: 0.00022438216547016054
step: 110, loss: 0.00030427734600380063
step: 120, loss: 0.10847436636686325
step: 130, loss: 0.0006251093000173569
step: 140, loss: 0.0011096361558884382
step: 150, loss: 0.004503467585891485
step: 160, loss: 0.0795804113149643
step: 170, loss: 0.01631866581737995
step: 180, loss: 0.0007847538799978793
step: 190, loss: 0.00810704380273819
step: 200, loss: 0.00023340796178672463
step: 210, loss: 0.0008772428263910115
step: 220, loss: 9.765536378836259e-05
step: 230, loss: 0.0006121547194197774
step: 240, loss: 0.004616624675691128
step: 250, loss: 0.0008017450454644859
step: 260, loss: 0.001819992670789361
step: 270, loss: 0.05866815149784088
step: 280, loss: 0.008714078925549984
step: 290, loss: 0.00025660410756245255
step: 300, loss: 0.02006078138947487
step: 310, loss: 0.0002855593338608742
step: 320, loss: 0.005089760757982731
step: 330, loss: 0.0023563187569379807
step: 340, loss: 0.006797105073928833
step: 350, loss: 0.004236483946442604
step: 360, loss: 0.0017374396556988358
step: 370, loss: 0.001075669890269637

```

```

step: 380, loss: 0.03347006440162659
step: 390, loss: 0.0055366153828799725
step: 400, loss: 0.009322794154286385
step: 410, loss: 0.000365587038686499
step: 420, loss: 0.0026035259943455458
step: 430, loss: 0.011473273858428001
step: 440, loss: 0.0004476552130654454
step: 460, loss: 0.0205231960862875
step: 470, loss: 0.015342304483056068
step: 480, loss: 0.006315311882644892
step: 490, loss: 0.006204657256603241
step: 500, loss: 0.0011911967303603888
step: 510, loss: 0.0014104587025940418
step: 520, loss: 0.00039861511322669685
step: 530, loss: 0.0028314010705798864
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.9276714885674288, f1=0.9265116279069768,
best_f1=0.9288354898336414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008167538326233625
step: 10, loss: 0.004392761737108231
step: 20, loss: 0.0006835705717094243
step: 30, loss: 0.00282928254455328
step: 40, loss: 0.005972210317850113
step: 50, loss: 0.00038999071693979204
step: 60, loss: 0.00013048957043793052
step: 70, loss: 0.0019586628768593073
step: 80, loss: 0.0003329449682496488
step: 90, loss: 0.005079175811260939
step: 100, loss: 0.0977855697274208
step: 110, loss: 0.007520898245275021
step: 120, loss: 0.0019338122801855206
step: 130, loss: 0.0007160475361160934
step: 140, loss: 0.009876308962702751
step: 150, loss: 0.004666339606046677
step: 160, loss: 0.0001734336547087878
step: 170, loss: 0.0006683908286504447
step: 180, loss: 0.001041053212247789
step: 190, loss: 0.00017660539015196264
step: 200, loss: 0.00010085326357511804
step: 210, loss: 0.000992062734439969
step: 220, loss: 0.00019334966782480478
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 230, loss: 0.0002657686418388039

```



```

step: 240, loss: 0.0002426494174869731
step: 250, loss: 0.0008397293859161437
step: 260, loss: 0.0019395491108298302
step: 270, loss: 0.010193562135100365
step: 280, loss: 0.0004905088571831584
step: 290, loss: 0.0010799919255077839
step: 300, loss: 0.0038492802996188402
step: 310, loss: 0.0022571550216525793
step: 320, loss: 0.010078747756779194
step: 330, loss: 0.007281574420630932
step: 340, loss: 0.016370924189686775
step: 350, loss: 0.015278110280632973
step: 360, loss: 0.00024335742637049407
step: 370, loss: 0.008219837211072445
step: 380, loss: 0.0005603085737675428
step: 390, loss: 0.1570506989955902
step: 400, loss: 0.0005789768183603883
step: 410, loss: 0.0007211731281131506
step: 420, loss: 0.05067579448223114
step: 430, loss: 0.0012250466970726848
step: 440, loss: 0.009312366135418415
step: 450, loss: 0.0019310768693685532
step: 460, loss: 0.0032523395493626595
step: 470, loss: 0.04517166689038277
step: 480, loss: 0.005616568028926849
step: 490, loss: 0.0043382043950259686
step: 500, loss: 0.0013127137208357453
step: 510, loss: 0.00044051173608750105
step: 520, loss: 0.0014224853366613388
step: 530, loss: 0.007559951860457659
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9241379310344828, f1=0.9229357798165138,
best_f1=0.9288354898336414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022521279752254486
step: 10, loss: 0.00755149545148015
step: 20, loss: 0.02498568221926689
step: 30, loss: 0.0026871338486671448
step: 40, loss: 0.016332531347870827
step: 50, loss: 0.04941166937351227
step: 60, loss: 0.08911855518817902
step: 70, loss: 0.0017510345205664635
step: 80, loss: 0.013559014536440372
step: 90, loss: 0.00012488231004681438

```

```
step: 100, loss: 0.004334860015660524
step: 110, loss: 0.007691734004765749
step: 120, loss: 0.00011365911632310599
step: 130, loss: 0.004849911667406559
step: 140, loss: 0.004089380148798227
step: 150, loss: 9.287580178352073e-05
step: 160, loss: 0.0003248709545005113
step: 170, loss: 0.005544343031942844
step: 180, loss: 0.0004035538004245609
step: 190, loss: 0.0016922417562454939
step: 200, loss: 0.1817130744457245
step: 210, loss: 0.006901120766997337
step: 220, loss: 0.000519425084348768
step: 230, loss: 0.000279231317108497
step: 240, loss: 0.004476028960198164
step: 250, loss: 0.00271128979511559
step: 260, loss: 0.0017128335312008858
step: 270, loss: 0.00011168060882482678
step: 280, loss: 0.05205777287483215
step: 290, loss: 0.008498304523527622
step: 300, loss: 0.0001730971271172166
step: 310, loss: 0.0003690744924824685
step: 320, loss: 0.045968879014253616
step: 330, loss: 0.0007213655626401305
step: 340, loss: 0.00883390847593546
step: 350, loss: 0.007773627061396837
step: 360, loss: 0.007674580905586481
step: 370, loss: 0.00023697152209933847
step: 380, loss: 0.0008761240751482546
step: 390, loss: 0.0002883868874050677
step: 400, loss: 0.0002468609600327909
step: 410, loss: 0.0010921660577878356
step: 420, loss: 0.0012314955238252878
step: 430, loss: 0.0004259438719600439
step: 440, loss: 0.0013391779502853751
step: 450, loss: 0.0014022127725183964
step: 460, loss: 0.017457690089941025
step: 470, loss: 0.0009203074732795358
step: 480, loss: 0.003160454798489809
step: 490, loss: 0.04066646471619606
step: 500, loss: 0.00014482683036476374
step: 510, loss: 0.009259807877242565
step: 520, loss: 0.0009808270260691643
step: 530, loss: 0.00013728336489293724
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9282385834109972, f1=0.9237248479176415,
```

```
best_f1=0.9288354898336414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004931556759402156
step: 10, loss: 0.04890850558876991
step: 20, loss: 0.021055741235613823
step: 30, loss: 0.01625463180243969
step: 40, loss: 0.0004674415395129472
step: 50, loss: 0.0005248987581580877
step: 60, loss: 0.0014807239640504122
step: 70, loss: 0.0007229701732285321
step: 80, loss: 0.0001485783141106367
step: 90, loss: 0.00010117578494828194
step: 100, loss: 7.879772601881996e-05
step: 110, loss: 0.0012728930450975895
step: 120, loss: 0.0015115354908630252
step: 130, loss: 0.0006763173150829971
step: 140, loss: 5.482901178766042e-05
step: 150, loss: 0.00016201770631596446
step: 160, loss: 0.00020111467165406793
step: 170, loss: 0.012629394419491291
step: 180, loss: 0.0007721410365775228
step: 190, loss: 0.0010009449906647205
step: 200, loss: 0.002307057147845626
step: 210, loss: 0.009809831157326698
step: 220, loss: 0.014920544810593128
step: 230, loss: 0.00018569121311884373
step: 240, loss: 0.00025568518321961164
step: 250, loss: 0.0027792525943368673
step: 260, loss: 0.009309777058660984
step: 270, loss: 0.00016558215429540724
step: 280, loss: 0.00022859667660668492
step: 290, loss: 0.0055792140774428844
step: 300, loss: 7.809663657099009e-05
step: 310, loss: 0.03378744050860405
step: 320, loss: 7.944284880068153e-05
step: 330, loss: 0.0033852483611553907
step: 340, loss: 0.000154172521433793
step: 350, loss: 0.0004321521846577525
step: 360, loss: 0.0002140999276889488
step: 370, loss: 0.00018014266970567405
step: 380, loss: 0.00018249389540869743
step: 390, loss: 9.658028284320608e-05
step: 400, loss: 0.030648667365312576
step: 410, loss: 0.0005321562639437616
step: 420, loss: 0.00022996067127678543
step: 430, loss: 0.02703203074634075
```

```

step: 440, loss: 0.02732272632420063
step: 450, loss: 0.0002988616470247507
step: 460, loss: 0.0038745037745684385
step: 470, loss: 0.008526810444891453
step: 480, loss: 0.0019283929141238332
step: 490, loss: 9.328432497568429e-05
step: 500, loss: 0.0020109356846660376
step: 510, loss: 0.00031319548725150526
step: 520, loss: 0.001963600981980562
step: 530, loss: 0.008631317876279354
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9234972677595629, f1=0.920562868815252,
best_f1=0.9288354898336414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015134003479033709
step: 10, loss: 0.0002367330016568303
step: 20, loss: 0.00022124066890683025
step: 30, loss: 0.00023853383027017117
step: 40, loss: 0.0004082688828930259
step: 50, loss: 0.0009386223391629755
step: 60, loss: 6.0067039157729596e-05
step: 70, loss: 0.09759359806776047
step: 80, loss: 0.00044843211071565747
step: 90, loss: 0.002019947161898017
step: 100, loss: 0.0016511313151568174
step: 110, loss: 0.0022626854479312897
step: 120, loss: 0.025521792471408844
step: 130, loss: 0.000412671419326216
step: 140, loss: 0.0021640651393681765
step: 150, loss: 0.00037731407792307436
step: 160, loss: 0.00039079601992852986
step: 170, loss: 0.00011066508886869997
step: 180, loss: 0.00012268272985238582
step: 190, loss: 0.0005904625286348164
step: 200, loss: 0.0004592994519043714
step: 210, loss: 7.386458310065791e-05
step: 220, loss: 0.0042440933175385
step: 230, loss: 0.0007140403031371534
step: 240, loss: 0.0408993661403656
step: 250, loss: 0.00017480722453910857
step: 260, loss: 0.0011710459366440773
step: 270, loss: 0.0005397769855335355
step: 280, loss: 9.08602523850277e-05
step: 290, loss: 0.0001829767134040594

```

```
step: 300, loss: 0.013952186331152916
step: 310, loss: 5.8219859056407586e-05
step: 320, loss: 0.0006736570503562689
step: 330, loss: 0.0001590853207744658
step: 340, loss: 0.00040948495734483004
step: 350, loss: 0.0001021712232613936
step: 360, loss: 0.01524156704545021
step: 370, loss: 0.0002461195399519056
step: 380, loss: 5.0585935241542757e-05
step: 390, loss: 0.015383601188659668
step: 400, loss: 0.03961785137653351
step: 410, loss: 0.00046299819950945675
step: 420, loss: 0.000761286006309092
step: 430, loss: 5.619449802907184e-05
step: 440, loss: 9.165146911982447e-05
step: 450, loss: 9.359233808936551e-05
step: 460, loss: 8.231796527979895e-05
step: 470, loss: 0.002900155959650874
step: 480, loss: 8.537405665265396e-05
step: 490, loss: 0.004898620303720236
step: 500, loss: 0.008017701096832752
step: 510, loss: 0.0020455194171518087
step: 520, loss: 0.0002555535756982863
step: 530, loss: 0.15160664916038513
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9259088817303267, f1=0.9279486002753558,
best_f1=0.9288354898336414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002945483720395714
step: 10, loss: 2.112956099153962e-05
step: 20, loss: 0.00025576326879672706
step: 30, loss: 6.327667506411672e-05
step: 40, loss: 5.392792809288949e-05
step: 50, loss: 4.353281474323012e-05
step: 60, loss: 0.0001417346065863967
step: 70, loss: 0.00013242047862149775
step: 80, loss: 0.003182257292792201
step: 90, loss: 0.0041696480475366116
step: 100, loss: 0.0002432225737720728
step: 110, loss: 0.0006694885087199509
step: 120, loss: 0.00790927279740572
step: 130, loss: 0.0005651421379297972
step: 140, loss: 0.00010189066961174831
step: 150, loss: 3.376036329427734e-05
```

```

step: 160, loss: 8.47887567942962e-05
step: 170, loss: 3.0472041544271633e-05
step: 180, loss: 0.0007598762167617679
step: 190, loss: 0.003054412780329585
step: 200, loss: 9.610851702746004e-05
step: 210, loss: 4.310301665100269e-05
step: 220, loss: 0.0012815567897632718
step: 230, loss: 8.489979518344626e-05
step: 240, loss: 0.00011287054803688079
step: 250, loss: 4.679220728576183e-05
step: 260, loss: 0.015374570153653622
step: 270, loss: 6.205132376635447e-05
step: 280, loss: 0.00045908731408417225
step: 290, loss: 6.394058436853811e-05
step: 300, loss: 0.00017879513325169683
step: 310, loss: 4.7522957174805924e-05
step: 320, loss: 0.0006318117375485599
step: 330, loss: 3.424550959607586e-05
step: 340, loss: 0.0008151286165229976
step: 350, loss: 6.083078551455401e-05
step: 360, loss: 0.00034105373197235167
step: 370, loss: 0.0014052852056920528
step: 380, loss: 0.00015844585141167045
step: 390, loss: 0.002058094134554267
step: 400, loss: 0.0024891183711588383
step: 410, loss: 0.0002691935806069523
step: 420, loss: 0.00024051083892118186
step: 430, loss: 9.065798803931102e-05
step: 440, loss: 0.00011000204540323466
step: 450, loss: 0.0001633006177144125
step: 460, loss: 0.0006008931086398661
step: 470, loss: 3.230092625017278e-05
step: 480, loss: 0.0032093648333102465
step: 490, loss: 0.02526511624455452
step: 500, loss: 0.00038108250009827316
step: 510, loss: 0.0003654375614132732
step: 520, loss: 0.019120359793305397
step: 530, loss: 5.016693103243597e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9224414869206058, f1=0.9228662711090826,
best_f1=0.9288354898336414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00046860656584613025
step: 10, loss: 0.0002567415649536997

```

step: 20, loss: 2.2578482457902282e-05  
step: 30, loss: 0.00013311395014170557  
step: 40, loss: 0.000347861903719604  
step: 50, loss: 6.531224789796397e-05  
step: 60, loss: 4.058554986841045e-05  
step: 70, loss: 0.00011025277490261942  
step: 80, loss: 0.3119255006313324  
step: 90, loss: 9.054982365341857e-05  
step: 100, loss: 0.019593236967921257  
step: 110, loss: 0.0005332007422111928  
step: 120, loss: 0.0003014520916622132  
step: 130, loss: 0.000280679581919685  
step: 140, loss: 0.00016818808217067271  
step: 150, loss: 0.00019950271234847605  
step: 160, loss: 0.003927238751202822  
step: 170, loss: 0.013080766424536705  
step: 180, loss: 0.00010285170719726011  
step: 190, loss: 0.0004252168000675738  
step: 200, loss: 0.007113078609108925  
step: 210, loss: 0.00013985300029162318  
step: 220, loss: 0.00011761716450564563  
step: 230, loss: 7.326862396439537e-05  
step: 240, loss: 6.402243889169767e-05  
step: 250, loss: 0.0003865232574753463  
step: 260, loss: 0.00014912016922608018  
step: 270, loss: 0.0007807198562659323  
step: 280, loss: 0.00012646916729863733  
step: 290, loss: 0.06263037770986557  
step: 300, loss: 0.001069962396286428  
step: 310, loss: 9.237596532329917e-05  
step: 320, loss: 0.010304946452379227  
step: 330, loss: 0.0012099252780899405  
step: 340, loss: 0.00048185494961217046  
step: 350, loss: 0.0017959775868803263  
step: 360, loss: 0.0032218429259955883  
step: 370, loss: 7.565187115687877e-05  
step: 380, loss: 4.4690743379760534e-05  
step: 390, loss: 0.002185279503464699  
step: 400, loss: 3.864515747409314e-05  
step: 410, loss: 8.212641114369035e-05  
step: 420, loss: 0.00010144953557755798  
step: 430, loss: 6.499258597614244e-05  
step: 440, loss: 6.091625618864782e-05  
step: 450, loss: 0.00013022981875110418  
step: 460, loss: 0.04065944626927376  
step: 470, loss: 0.0009142294875346124  
step: 480, loss: 0.00010783393372548744  
step: 490, loss: 0.00042672481504268944

```

step: 500, loss: 0.00014967862807679921
step: 510, loss: 0.0001689033815637231
step: 520, loss: 0.0001447843824280426
step: 530, loss: 0.0008766023674979806
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9255514705882354, f1=0.9275229357798166,
best_f1=0.9288354898336414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011828827700810507
step: 10, loss: 9.764882270246744e-05
step: 20, loss: 0.000538051943294704
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 30, loss: 0.0009198060142807662
step: 40, loss: 7.164091221056879e-05
step: 50, loss: 0.0007673365180380642
step: 60, loss: 0.00019403939950279891
step: 70, loss: 0.003074724692851305
step: 80, loss: 0.00019886587688233703
step: 90, loss: 0.00013861800835002214
step: 100, loss: 0.00010636037040967494
step: 110, loss: 8.214027184294537e-05
step: 120, loss: 6.574431608896703e-05
step: 130, loss: 9.606347884982824e-05
step: 140, loss: 6.029766154824756e-05
step: 150, loss: 0.0893712043762207
step: 160, loss: 0.00019990169676020741
step: 170, loss: 4.621059633791447e-05
step: 180, loss: 0.0018699561478570104
step: 190, loss: 0.0005575491231866181
step: 200, loss: 9.9722616141662e-05
step: 210, loss: 0.04691728949546814
step: 220, loss: 0.00010023711365647614
step: 230, loss: 0.00045074912486597896
step: 240, loss: 0.00027956353733316064
step: 250, loss: 8.549694757675752e-05
step: 260, loss: 0.00022876312141306698
step: 270, loss: 0.26609745621681213
step: 280, loss: 0.00029202154837548733
step: 290, loss: 0.001217341748997569
step: 300, loss: 0.0005246620276011527
step: 310, loss: 0.00018722984532359987
step: 320, loss: 0.0001361541508231312
step: 330, loss: 0.00018476835975889117
step: 340, loss: 0.0003538944583851844

```



```

step: 350, loss: 0.0067829047329723835
step: 360, loss: 0.00012478252756409347
step: 370, loss: 0.00011074247595388442
step: 380, loss: 0.00012601504568010569
step: 390, loss: 0.00012072856770828366
step: 400, loss: 0.00020778765610884875
step: 410, loss: 0.0005668496014550328
step: 420, loss: 0.0009066187194548547
step: 430, loss: 0.0002162058954127133
step: 440, loss: 0.001018830225802958
step: 450, loss: 0.0011576018296182156
step: 460, loss: 0.0011781379580497742
step: 470, loss: 0.00038013909943401814
step: 480, loss: 7.732827361905947e-05
step: 490, loss: 0.0024263823870569468
step: 500, loss: 0.0005611918168142438
step: 510, loss: 0.0001500382204540074
step: 520, loss: 0.019536254927515984
step: 530, loss: 0.00016152617172338068
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9188445667125174, f1=0.9258406264394289,
best_f1=0.9288354898336414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.646751328138635e-05
step: 10, loss: 0.00011838630598504096
step: 20, loss: 0.014030861668288708
step: 30, loss: 0.0063294535502791405
step: 40, loss: 0.0008836037595756352
step: 50, loss: 0.00013399694580584764
step: 60, loss: 0.0005240619066171348
step: 70, loss: 6.68601569486782e-05
step: 80, loss: 5.041239637648687e-05
step: 90, loss: 0.00018285516125615686
step: 100, loss: 5.340348434401676e-05
step: 110, loss: 7.405552605632693e-05
step: 120, loss: 6.020904038450681e-05
step: 130, loss: 0.005042994860559702
step: 140, loss: 0.001359636546112597
step: 150, loss: 0.00012530699314083904
step: 160, loss: 7.780407031532377e-05
step: 170, loss: 0.00017426109116058797
step: 180, loss: 9.375846275361255e-05
step: 190, loss: 0.0002137581177521497
step: 200, loss: 0.0055654835887253284

```

```

step: 210, loss: 0.0011342833749949932
step: 220, loss: 6.362282147165388e-05
step: 230, loss: 8.004721166798845e-05
step: 240, loss: 0.00010598354128887877
step: 250, loss: 0.08439943939447403
step: 260, loss: 0.00010027505049947649
step: 270, loss: 6.670570292044431e-05
step: 280, loss: 4.888078183284961e-05
step: 290, loss: 0.002177101094275713
step: 300, loss: 0.00012551850522868335
step: 310, loss: 0.006374312099069357
step: 320, loss: 4.191916013951413e-05
step: 330, loss: 9.260346996597946e-05
step: 340, loss: 9.64958526310511e-05
step: 350, loss: 0.02141410857439041
step: 360, loss: 5.303844955051318e-05
step: 370, loss: 0.00011271281982772052
step: 380, loss: 0.001715178368613124
step: 390, loss: 0.00056216842494905
step: 400, loss: 4.2697654862422496e-05
step: 410, loss: 5.113138831802644e-05
step: 420, loss: 4.5875283831264824e-05
step: 430, loss: 0.0010643013520166278
step: 440, loss: 0.00016815192066133022
step: 450, loss: 0.00010211457993136719
step: 460, loss: 3.6517380067380145e-05
step: 470, loss: 0.0012660648208111525
step: 480, loss: 0.00040525899385102093
step: 490, loss: 7.285321044037119e-05
step: 500, loss: 0.00016802820027805865
step: 510, loss: 5.227417204878293e-05
step: 520, loss: 6.588854012079537e-05
step: 530, loss: 0.0001464543747715652
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9208566108007449, f1=0.9284386617100371,
best_f1=0.9288354898336414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0074743954464793205
step: 10, loss: 0.00014986972382757813
step: 20, loss: 3.0427432648139074e-05
step: 30, loss: 8.485355647280812e-05
step: 40, loss: 4.322112727095373e-05
step: 50, loss: 0.0005342158838175237
step: 60, loss: 5.247507942840457e-05

```

step: 70, loss: 3.210000795661472e-05  
step: 80, loss: 0.00028217246290296316  
step: 90, loss: 2.6884779799729586e-05  
step: 100, loss: 4.9092861445387825e-05  
step: 110, loss: 3.318766539450735e-05  
step: 120, loss: 3.5064833355136216e-05  
step: 130, loss: 6.713929178658873e-05  
step: 140, loss: 0.01131188403815031  
step: 150, loss: 5.730461634811945e-05  
step: 160, loss: 0.0005420449888333678  
step: 170, loss: 8.862609684001654e-05  
step: 180, loss: 5.0451366405468434e-05  
step: 190, loss: 7.677274697925895e-05  
step: 200, loss: 5.2499639423331246e-05  
step: 210, loss: 7.908744009910151e-05  
step: 220, loss: 4.908815026283264e-05  
step: 230, loss: 4.910132702207193e-05  
step: 240, loss: 4.4886870455229655e-05  
step: 250, loss: 4.975797492079437e-05  
step: 260, loss: 2.6702262402977794e-05  
step: 270, loss: 0.005713874474167824  
step: 280, loss: 5.286180748953484e-05  
step: 290, loss: 5.598288043984212e-05  
step: 300, loss: 5.839706500410102e-05  
step: 310, loss: 4.89897902298253e-05  
step: 320, loss: 0.00012971276009920985  
step: 330, loss: 4.9655242037260905e-05  
step: 340, loss: 9.737018262967467e-05  
step: 350, loss: 0.00010536804620642215  
step: 360, loss: 0.0035290298983454704  
step: 370, loss: 9.335883805761114e-05  
step: 380, loss: 3.5034401662414894e-05  
step: 390, loss: 4.836914376937784e-05  
step: 400, loss: 3.0840841645840555e-05  
step: 410, loss: 6.906605267431587e-05  
step: 420, loss: 6.983029743423685e-05  
step: 430, loss: 0.00010274178202962503  
step: 440, loss: 2.3960616090334952e-05  
step: 450, loss: 2.95631598419277e-05  
step: 460, loss: 0.0003674962208606303  
step: 470, loss: 3.225610635126941e-05  
step: 480, loss: 3.839895725832321e-05  
step: 490, loss: 7.739668217254803e-05  
step: 500, loss: 3.651397491921671e-05  
step: 510, loss: 0.00020189899078104645  
step: 520, loss: 6.253230822039768e-05  
step: 530, loss: 2.5171377274091356e-05  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning:

```
An input tensor was not cuda.  
  warnings.warn("An input tensor was not cuda.")  
epoch 14: dev_f1=0.9225806451612903, f1=0.927816091954023,  
best_f1=0.9288354898336414  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
  warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.00018725790141616017  
step: 10, loss: 3.825331077678129e-05  
step: 20, loss: 6.676763587165624e-05  
step: 30, loss: 4.867448296863586e-05  
step: 40, loss: 0.00014765300147701055  
step: 50, loss: 7.401351467706263e-05  
step: 60, loss: 6.223251693882048e-05  
step: 70, loss: 7.656320667592809e-05  
step: 80, loss: 2.8736048989230767e-05  
step: 90, loss: 7.342418393818662e-05  
step: 100, loss: 2.5525117962388322e-05  
step: 110, loss: 0.016067320480942726  
step: 120, loss: 0.00012795760994777083  
step: 130, loss: 2.1375317373895086e-05  
step: 140, loss: 1.956494634214323e-05  
step: 150, loss: 4.822497794521041e-05  
step: 160, loss: 0.0001295414549531415  
step: 170, loss: 0.00010464766091899946  
step: 180, loss: 6.482742173830047e-05  
step: 190, loss: 0.00011883532715728506  
step: 200, loss: 3.405870302231051e-05  
step: 210, loss: 6.513814878417179e-05  
step: 220, loss: 2.9995018849149346e-05  
step: 230, loss: 0.0002548295015003532  
step: 240, loss: 3.633143933257088e-05  
step: 250, loss: 3.292673864052631e-05  
step: 260, loss: 5.122414222569205e-05  
step: 270, loss: 0.00010756815754575655  
step: 280, loss: 5.638975926558487e-05  
step: 290, loss: 9.18531441129744e-05  
step: 300, loss: 0.0007352014654316008  
step: 310, loss: 6.567737727891654e-05  
step: 320, loss: 4.8361449444200844e-05  
step: 330, loss: 0.00010573264444246888  
step: 340, loss: 6.219738133950159e-05  
step: 350, loss: 0.00013479258632287383  
step: 360, loss: 7.091284351190552e-05  
step: 370, loss: 2.5260533220716752e-05  
step: 380, loss: 4.672547220252454e-05  
step: 390, loss: 3.476253550616093e-05  
step: 400, loss: 2.2686672309646383e-05
```

```

step: 410, loss: 4.882643770542927e-05
step: 420, loss: 5.469836469274014e-05
step: 430, loss: 2.6173232981818728e-05
step: 440, loss: 7.595994975417852e-05
step: 450, loss: 0.0001829243847168982
step: 460, loss: 2.3390628484776244e-05
step: 470, loss: 2.4850813133525662e-05
step: 480, loss: 0.00019616902864072472
step: 490, loss: 3.0576251447200775e-05
step: 500, loss: 5.2498289733193815e-05
step: 510, loss: 0.00011715363507391885
step: 520, loss: 3.005016696988605e-05
step: 530, loss: 2.5081808416871354e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9208699676075892, f1=0.9251386321626617,
best_f1=0.9288354898336414
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

5742it [00:14, 400.40it/s]
load_f1 = 0.9349930843706776
real_f1 = 0.9335180055401663
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:11, 399.02it/s]

```

## 2.1.4 Amazon-Google - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/Amazon-Google \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm distilbert \
  --fp16 \
  --save_model

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Amazon-Google.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/Amazon-Google \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm distilbert \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False

```

```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                   : True
opt_level                 : 02
cast_model_type           : torch.float16
patch_torch_functions     : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.8376253843307495
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.061901409178972244
step: 20, loss: 0.3784135580062866
step: 30, loss: 0.3762974739074707
step: 40, loss: 0.5026069283485413
step: 50, loss: 0.29514601826667786
step: 60, loss: 0.3615758419036865
step: 70, loss: 0.22282052040100098
step: 80, loss: 0.27890026569366455
step: 90, loss: 0.4078083038330078
step: 100, loss: 0.17996463179588318
step: 110, loss: 0.3094843029975891
step: 120, loss: 0.21902351081371307
step: 130, loss: 0.27711689472198486
step: 140, loss: 0.2522067129611969
step: 150, loss: 0.27141833305358887
step: 160, loss: 0.20531830191612244
step: 170, loss: 0.11413554847240448
step: 180, loss: 0.21554933488368988

```

```

step: 190, loss: 0.2299884706735611
step: 200, loss: 0.17916886508464813
step: 210, loss: 0.48345044255256653
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.5284552845528455, f1=0.530214424951267,
best_f1=0.530214424951267
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11060486733913422
step: 10, loss: 0.07949305325746536
step: 20, loss: 0.30779722332954407
step: 30, loss: 0.15375851094722748
step: 40, loss: 0.1189803034067154
step: 50, loss: 0.21055570244789124
step: 60, loss: 0.06594382226467133
step: 70, loss: 0.14837661385536194
step: 80, loss: 0.18305020034313202
step: 90, loss: 0.22953647375106812
step: 100, loss: 0.09149108082056046
step: 110, loss: 0.09397146105766296
step: 120, loss: 0.27938181161880493
step: 130, loss: 0.18459026515483856
step: 140, loss: 0.22829371690750122
step: 150, loss: 0.2029259204864502
step: 160, loss: 0.14308354258537292
step: 170, loss: 0.2256779968738556
step: 180, loss: 0.31439656019210815
step: 190, loss: 0.16303041577339172
step: 200, loss: 0.23074494302272797
step: 210, loss: 0.28391024470329285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.5880341880341882, f1=0.6023688663282573,
best_f1=0.6023688663282573
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26455673575401306
step: 10, loss: 0.17930682003498077
step: 20, loss: 0.2792327404022217
step: 30, loss: 0.032110027968883514
step: 40, loss: 0.20582300424575806
step: 50, loss: 0.10732422024011612
step: 60, loss: 0.1982710063457489
step: 70, loss: 0.09560135006904602
step: 80, loss: 0.04137900471687317
step: 90, loss: 0.14851711690425873
step: 100, loss: 0.04160679504275322
step: 110, loss: 0.10157215595245361
step: 120, loss: 0.2700756788253784
step: 130, loss: 0.10322149842977524
step: 140, loss: 0.23135335743427277
step: 150, loss: 0.1989177167415619
step: 160, loss: 0.04559176787734032
step: 170, loss: 0.11518287658691406
step: 180, loss: 0.022503551095724106
step: 190, loss: 0.27714115381240845
step: 200, loss: 0.15233245491981506
step: 210, loss: 0.21395643055438995
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.5860927152317881, f1=0.5927209705372616,
best_f1=0.6023688663282573
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06656055152416229
step: 10, loss: 0.060229018330574036
step: 20, loss: 0.14841634035110474
step: 30, loss: 0.06528966128826141
step: 40, loss: 0.05566629767417908
step: 50, loss: 0.3218264877796173
step: 60, loss: 0.03627605363726616
step: 70, loss: 0.11999905109405518
step: 80, loss: 0.18595002591609955
step: 90, loss: 0.004276525229215622
step: 100, loss: 0.09131759405136108
step: 110, loss: 0.18393179774284363
step: 120, loss: 0.050940223038196564
step: 130, loss: 0.27148082852363586
step: 140, loss: 0.1547771841287613
step: 150, loss: 0.1120658740401268
step: 160, loss: 0.06873100996017456

```

```

step: 170, loss: 0.028330113738775253
step: 180, loss: 0.11479097604751587
step: 190, loss: 0.08389890938997269
step: 200, loss: 0.09108783304691315
step: 210, loss: 0.013936229050159454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.5840707964601769, f1=0.5669642857142856,
best_f1=0.6023688663282573
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06373542547225952
step: 10, loss: 0.0955292358994484
step: 20, loss: 0.01655951701104641
step: 30, loss: 0.13165812194347382
step: 40, loss: 0.12831096351146698
step: 50, loss: 0.09845130145549774
step: 60, loss: 0.009696329943835735
step: 70, loss: 0.04013989493250847
step: 80, loss: 0.026778601109981537
step: 90, loss: 0.09877508133649826
step: 100, loss: 0.07144832611083984
step: 110, loss: 0.006220249459147453
step: 120, loss: 0.006164159160107374
step: 130, loss: 0.1193370670080185
step: 140, loss: 0.0733642652630806
step: 150, loss: 0.03998799994587898
step: 160, loss: 0.008369773626327515
step: 170, loss: 0.04722139239311218
step: 180, loss: 0.11638817936182022
step: 190, loss: 0.16687673330307007
step: 200, loss: 0.0790741965174675
step: 210, loss: 0.07862625271081924
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.6056910569105691, f1=0.5635593220338982,
best_f1=0.5635593220338982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06426317244768143

```

```

step: 10, loss: 0.07412722706794739
step: 20, loss: 0.020533738657832146
step: 30, loss: 0.12801513075828552
step: 40, loss: 0.034808821976184845
step: 50, loss: 0.05016127973794937
step: 60, loss: 0.18601974844932556
step: 70, loss: 0.001595427980646491
step: 80, loss: 0.07126922905445099
step: 90, loss: 0.11567935347557068
step: 100, loss: 0.029189616441726685
step: 110, loss: 0.042621366679668427
step: 120, loss: 0.022514892742037773
step: 130, loss: 0.0269310362637043
step: 140, loss: 0.040593620389699936
step: 150, loss: 0.01497785933315754
step: 160, loss: 0.0769072026014328
step: 170, loss: 0.0294471625238657
step: 180, loss: 0.0021561570465564728
step: 190, loss: 0.05436651408672333
step: 200, loss: 0.007846180349588394
step: 210, loss: 0.02191273681819439
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.5907172995780591, f1=0.5543237250554324,
best_f1=0.5635593220338982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011471966281533241
step: 10, loss: 0.005814142990857363
step: 20, loss: 0.01726415567100048
step: 30, loss: 0.008614332415163517
step: 40, loss: 0.1241883859038353
step: 50, loss: 0.02542215771973133
step: 60, loss: 0.36206626892089844
step: 70, loss: 0.004023829009383917
step: 80, loss: 0.0018792699556797743
step: 90, loss: 0.01575653627514839
step: 100, loss: 0.010438580997288227
step: 110, loss: 0.06956997513771057
step: 120, loss: 0.06007693335413933
step: 130, loss: 0.018801933154463768
step: 140, loss: 0.004767714533954859
step: 150, loss: 0.061939261853694916
step: 160, loss: 0.030241886153817177
step: 170, loss: 0.06661226600408554
step: 180, loss: 0.0024773620534688234

```

```

step: 190, loss: 0.025002969428896904
step: 200, loss: 0.006949182134121656
step: 210, loss: 0.01905723102390766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.5755102040816326, f1=0.5874730021598272,
best_f1=0.5635593220338982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.049382880330085754
step: 10, loss: 0.005252340342849493
step: 20, loss: 0.04104415699839592
step: 30, loss: 0.014280476607382298
step: 40, loss: 0.00909211952239275
step: 50, loss: 0.0033333159517496824
step: 60, loss: 0.1270262598991394
step: 70, loss: 0.0038734967820346355
step: 80, loss: 0.006696522235870361
step: 90, loss: 0.009491453878581524
step: 100, loss: 0.002193504013121128
step: 110, loss: 0.004785217344760895
step: 120, loss: 0.002064506057649851
step: 130, loss: 0.0057804277166724205
step: 140, loss: 0.04185056313872337
step: 150, loss: 0.004608507268130779
step: 160, loss: 0.01449884194880724
step: 170, loss: 0.004789073020219803
step: 180, loss: 0.048969559371471405
step: 190, loss: 0.04275822639465332
step: 200, loss: 0.017632734030485153
step: 210, loss: 0.08315318077802658
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.5901639344262296, f1=0.5769230769230769,
best_f1=0.5635593220338982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.027217980474233627
step: 10, loss: 0.01047193631529808
step: 20, loss: 0.01871372014284134
step: 30, loss: 0.05178769305348396
step: 40, loss: 0.08172901719808578
step: 50, loss: 0.036696452647447586
step: 60, loss: 0.003423749702051282

```

```

step: 70, loss: 0.00533897103741765
step: 80, loss: 0.028403569012880325
step: 90, loss: 0.12184469401836395
step: 100, loss: 0.0010051354765892029
step: 110, loss: 0.012213676236569881
step: 120, loss: 0.0015966454520821571
step: 130, loss: 0.0060348245315253735
step: 140, loss: 0.06934812664985657
step: 150, loss: 0.04233839735388756
step: 160, loss: 0.01602206937968731
step: 170, loss: 0.008356329053640366
step: 180, loss: 0.007647150196135044
step: 190, loss: 0.02854614518582821
step: 200, loss: 0.03435765579342842
step: 210, loss: 0.009419983252882957
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.5720164609053499, f1=0.5743801652892562,
best_f1=0.5635593220338982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026629972271621227
step: 10, loss: 0.00044609446194954216
step: 20, loss: 0.02849246934056282
step: 30, loss: 0.012109307572245598
step: 40, loss: 0.08653198182582855
step: 50, loss: 0.0024329624138772488
step: 60, loss: 0.00022538223129231483
step: 70, loss: 0.001602771575562656
step: 80, loss: 0.0002472419000696391
step: 90, loss: 0.0045324829407036304
step: 100, loss: 0.00017964481958188117
step: 110, loss: 0.002207403304055333
step: 120, loss: 0.0027047479525208473
step: 130, loss: 0.0008428182918578386
step: 140, loss: 0.001236427458934486
step: 150, loss: 0.004387567285448313
step: 160, loss: 0.013930769637227058
step: 170, loss: 0.025400256738066673
step: 180, loss: 0.0007615337963216007
step: 190, loss: 0.001694022910669446
step: 200, loss: 0.03557126596570015
step: 210, loss: 0.007667890749871731
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

epoch 10: dev_f1=0.579476861167002, f1=0.5619834710743802,
best_f1=0.5635593220338982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014381424523890018
step: 10, loss: 0.01745898276567459
step: 20, loss: 0.0005640284507535398
step: 30, loss: 0.057641174644231796
step: 40, loss: 0.006635839119553566
step: 50, loss: 0.0010131526505574584
step: 60, loss: 0.02814166620373726
step: 70, loss: 0.0015788362361490726
step: 80, loss: 0.10487734526395798
step: 90, loss: 0.012746266089379787
step: 100, loss: 0.009155649691820145
step: 110, loss: 0.004305624403059483
step: 120, loss: 0.0017738096648827195
step: 130, loss: 0.04077215865254402
step: 140, loss: 0.016671253368258476
step: 150, loss: 0.05890533700585365
step: 160, loss: 0.0005509629263542593
step: 170, loss: 0.11603403836488724
step: 180, loss: 0.007967192679643631
step: 190, loss: 0.04229258373379707
step: 200, loss: 0.01548620406538248
step: 210, loss: 0.003008089726790786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.5864661654135338, f1=0.5868263473053893,
best_f1=0.5635593220338982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00036049270420335233
step: 10, loss: 0.0017609490314498544
step: 20, loss: 0.0007099923677742481
step: 30, loss: 0.000284995068795979
step: 40, loss: 0.02742685377597809
step: 50, loss: 0.025451235473155975
step: 60, loss: 0.003290207590907812
step: 70, loss: 0.00024479886633343995
step: 80, loss: 0.0003408722404856235
step: 90, loss: 0.00035816416493617
step: 100, loss: 0.0029280688613653183
step: 110, loss: 0.0005146154435351491
step: 120, loss: 0.003515060292556882

```

```

step: 130, loss: 0.021802501752972603
step: 140, loss: 0.0011722627095878124
step: 150, loss: 0.011655454523861408
step: 160, loss: 0.0006204845267347991
step: 170, loss: 0.01204589381814003
step: 180, loss: 0.0022756517864763737
step: 190, loss: 0.010604691691696644
step: 200, loss: 0.1348857283592224
step: 210, loss: 0.12866279482841492
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.593186372745491, f1=0.5743801652892562,
best_f1=0.5635593220338982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00024424109142273664
step: 10, loss: 0.006894251331686974
step: 20, loss: 0.000337052479153499
step: 30, loss: 0.01977875456213951
step: 40, loss: 0.00033179958700202405
step: 50, loss: 0.00030530410003848374
step: 60, loss: 0.0022815184202045202
step: 70, loss: 0.02605544775724411
step: 80, loss: 0.003981141373515129
step: 90, loss: 0.028004884719848633
step: 100, loss: 0.00767943263053894
step: 110, loss: 0.0007675565429963171
step: 120, loss: 0.00028005396598018706
step: 130, loss: 0.0002168371865991503
step: 140, loss: 0.011373150162398815
step: 150, loss: 0.0004911042051389813
step: 160, loss: 0.001204965403303504
step: 170, loss: 0.001713807345367968
step: 180, loss: 0.09930839389562607
step: 190, loss: 0.00017966334416996688
step: 200, loss: 0.0002371323644183576
step: 210, loss: 0.0004409267276059836
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.5806451612903226, f1=0.5738045738045738,
best_f1=0.5635593220338982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017555195838212967

```

```

step: 10, loss: 0.0002429726446280256
step: 20, loss: 0.0007656249799765646
step: 30, loss: 0.000391207926440984
step: 40, loss: 0.003945067524909973
step: 50, loss: 0.0009094974957406521
step: 60, loss: 0.003760540159419179
step: 70, loss: 0.002064401051029563
step: 80, loss: 0.05830944702029228
step: 90, loss: 0.0007518883212469518
step: 100, loss: 0.0008438549702987075
step: 110, loss: 0.00037478411104530096
step: 120, loss: 0.00019789765065070242
step: 130, loss: 0.0001609104947419837
step: 140, loss: 0.00024214516452047974
step: 150, loss: 0.0005098524270579219
step: 160, loss: 0.0003568599931895733
step: 170, loss: 0.0001866481761680916
step: 180, loss: 0.0007626009173691273
step: 190, loss: 0.0002209292579209432
step: 200, loss: 0.0001309887011302635
step: 210, loss: 0.000757779402192682
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.5880077369439072, f1=0.5766871165644172,
best_f1=0.5635593220338982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01272359024733305
step: 10, loss: 0.001928345300257206
step: 20, loss: 0.0001707689225440845
step: 30, loss: 0.0008084952714852989
step: 40, loss: 0.00015789475583005697
step: 50, loss: 0.0002735721645876765
step: 60, loss: 0.001239356235601008
step: 70, loss: 0.00526912696659565
step: 80, loss: 0.0003586901293601841
step: 90, loss: 0.002702889731153846
step: 100, loss: 0.0011551727075129747
step: 110, loss: 0.00011701555195031688
step: 120, loss: 0.0012991838157176971
step: 130, loss: 0.0021700235083699226
step: 140, loss: 0.000657066993881017
step: 150, loss: 0.00029318095766939223
step: 160, loss: 0.0010859068715944886
step: 170, loss: 0.026246244087815285
step: 180, loss: 0.00343522010371089

```



```

step: 190, loss: 0.0005715562147088349
step: 200, loss: 0.0772310197353363
step: 210, loss: 0.02183850295841694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.5875486381322957, f1=0.5737704918032788,
best_f1=0.5635593220338982
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2293it [00:03, 653.05it/s]
load_f1 = 0.6160164271047228
real_f1 = 0.6044624746450304
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:10, 404.67it/s]

```

### 2.1.5 Walmart-Amazon - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
--task Structured/Walmart-Amazon \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm distilbert \
--fp16 \
--save_model

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Walmart-Amazon.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
--task Structured/Walmart-Amazon \
--input_path input/input_small.jsonl \
--output_path output/output_small.jsonl \
--lm distilbert \
--use_gpu \
--fp16 \
--checkpoint_path checkpoints/
```

Downloading: 100% 442/442 [00:00<00:00, 373kB/s]

Downloading: 100% 232k/232k [00:00<00:00, 257kB/s]

Downloading: 100% 268M/268M [00:05<00:00, 48.8MB/s]

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

enabled : True

opt\_level : 02

cast\_model\_type : torch.float16

patch\_torch\_functions : False

keep\_batchnorm\_fp32 : True

master\_weights : True

loss\_scale : dynamic

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

enabled : True

```

opt_level          : 02
cast_model_type    : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights     : True
loss_scale         : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.8631476163864136
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.15778599679470062
step: 20, loss: 0.155282124876976
step: 30, loss: 0.5144890546798706
step: 40, loss: 0.25378385186195374
step: 50, loss: 0.31217721104621887
step: 60, loss: 0.36030614376068115
step: 70, loss: 0.18065176904201508
step: 80, loss: 0.5272958278656006
step: 90, loss: 0.2306559681892395
step: 100, loss: 0.2205209732055664
step: 110, loss: 0.23886151611804962
step: 120, loss: 0.42526328563690186
step: 130, loss: 0.3489666283130646
step: 140, loss: 0.3290039002895355
step: 150, loss: 0.26765426993370056
step: 160, loss: 0.21536333858966827
step: 170, loss: 0.40876492857933044
step: 180, loss: 0.28496024012565613
step: 190, loss: 0.15831685066223145
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or

```

```

loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.37068965517241376, f1=0.38362068965517243,
best_f1=0.38362068965517243
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31150001287460327
step: 10, loss: 0.047406554222106934
step: 20, loss: 0.18013368546962738
step: 30, loss: 0.17602966725826263
step: 40, loss: 0.4906077980995178
step: 50, loss: 0.2786729335784912
step: 60, loss: 0.1680448055267334
step: 70, loss: 0.2336992472410202
step: 80, loss: 0.10713095963001251
step: 90, loss: 0.12542152404785156
step: 100, loss: 0.2801400423049927
step: 110, loss: 0.13899832963943481
step: 120, loss: 0.3260059952735901
step: 130, loss: 0.23339983820915222
step: 140, loss: 0.26732635498046875
step: 150, loss: 0.05662167817354202
step: 160, loss: 0.12306304275989532
step: 170, loss: 0.3041684329509735
step: 180, loss: 0.17113138735294342
step: 190, loss: 0.20091219246387482
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.6774193548387096, f1=0.6814404432132963,
best_f1=0.6814404432132963
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1474648118019104
step: 10, loss: 0.2375137060880661
step: 20, loss: 0.04162169247865677
step: 30, loss: 0.027241485193371773
step: 40, loss: 0.02865813672542572
step: 50, loss: 0.17535310983657837
step: 60, loss: 0.09536435455083847
step: 70, loss: 0.10306625813245773
step: 80, loss: 0.10132929682731628

```

```

step: 90, loss: 0.20426703989505768
step: 100, loss: 0.06540381163358688
step: 110, loss: 0.17636221647262573
step: 120, loss: 0.012787170708179474
step: 130, loss: 0.04447212442755699
step: 140, loss: 0.1870233118534088
step: 150, loss: 0.0930280014872551
step: 160, loss: 0.12259366363286972
step: 170, loss: 0.06549486517906189
step: 180, loss: 0.07564670592546463
step: 190, loss: 0.1583467721939087
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.7247191011235956, f1=0.7449856733524355,
best_f1=0.7449856733524355
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03252561017870903
step: 10, loss: 0.031142299994826317
step: 20, loss: 0.06062079593539238
step: 30, loss: 0.07372695207595825
step: 40, loss: 0.006952584255486727
step: 50, loss: 0.02688789740204811
step: 60, loss: 0.056491609662771225
step: 70, loss: 0.007189827039837837
step: 80, loss: 0.0365331694483757
step: 90, loss: 0.02446473389863968
step: 100, loss: 0.016053596511483192
step: 110, loss: 0.0820852741599083
step: 120, loss: 0.03935001417994499
step: 130, loss: 0.4311694800853729
step: 140, loss: 0.07165408134460449
step: 150, loss: 0.07884325087070465
step: 160, loss: 0.06940142065286636
step: 170, loss: 0.010451861657202244
step: 180, loss: 0.06965596973896027
step: 190, loss: 0.06203407794237137
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or

```

```

loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.7292225201072386, f1=0.743801652892562,
best_f1=0.743801652892562
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04550010710954666
step: 10, loss: 0.007558650337159634
step: 20, loss: 0.03487584367394447
step: 30, loss: 0.01732426881790161
step: 40, loss: 0.02983095869421959
step: 50, loss: 0.013548186980187893
step: 60, loss: 0.061976008117198944
step: 70, loss: 0.042256057262420654
step: 80, loss: 0.060228414833545685
step: 90, loss: 0.06522847712039948
step: 100, loss: 0.02871328592300415
step: 110, loss: 0.008076585829257965
step: 120, loss: 0.002543156500905752
step: 130, loss: 0.006084659602493048
step: 140, loss: 0.005587242543697357
step: 150, loss: 0.016255775466561317
step: 160, loss: 0.008055290207266808
step: 170, loss: 0.022079724818468094
step: 180, loss: 0.08114727586507797
step: 190, loss: 0.17327576875686646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.7336956521739131, f1=0.7374301675977653,
best_f1=0.7374301675977653
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022185999900102615
step: 10, loss: 0.0017828033305704594
step: 20, loss: 0.034126151353120804
step: 30, loss: 0.005033095832914114
step: 40, loss: 0.030655808746814728
step: 50, loss: 0.13865691423416138
step: 60, loss: 0.00038786345976404846
step: 70, loss: 0.05302300676703453
step: 80, loss: 0.0455089695751667

```

```

step: 90, loss: 0.0025861491449177265
step: 100, loss: 0.019096272066235542
step: 110, loss: 0.02115963213145733
step: 120, loss: 0.024637490510940552
step: 130, loss: 0.0021130028180778027
step: 140, loss: 0.0018642973154783249
step: 150, loss: 0.02659563161432743
step: 160, loss: 0.00267317658290267
step: 170, loss: 0.044997021555900574
step: 180, loss: 0.002769424580037594
step: 190, loss: 0.016683926805853844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.7277486910994764, f1=0.7311827956989247,
best_f1=0.7374301675977653
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014075905783101916
step: 10, loss: 0.00838404055684805
step: 20, loss: 0.002636388875544071
step: 30, loss: 0.008160828612744808
step: 40, loss: 0.0016545649850741029
step: 50, loss: 0.05994168668985367
step: 60, loss: 0.0008226595818996429
step: 70, loss: 0.0008546877652406693
step: 80, loss: 0.05911964923143387
step: 90, loss: 0.001654914696700871
step: 100, loss: 0.02767769619822502
step: 110, loss: 0.0034609807189553976
step: 120, loss: 0.0032049515284597874
step: 130, loss: 0.04279407113790512
step: 140, loss: 0.0019837357103824615
step: 150, loss: 0.004245641641318798
step: 160, loss: 0.0007435794686898589
step: 170, loss: 0.003915438428521156
step: 180, loss: 0.008151288144290447
step: 190, loss: 0.0008012704784050584
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.7135678391959798, f1=0.760204081632653,
best_f1=0.7374301675977653
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000359549856511876

```

```

step: 10, loss: 0.0010025780647993088
step: 20, loss: 0.0008331550634466112
step: 30, loss: 0.0029197982512414455
step: 40, loss: 0.0031706171575933695
step: 50, loss: 0.0009271536837331951
step: 60, loss: 0.0010286916512995958
step: 70, loss: 0.008196474984288216
step: 80, loss: 0.023123404011130333
step: 90, loss: 0.0009789093164727092
step: 100, loss: 0.0009705307311378419
step: 110, loss: 0.004233557730913162
step: 120, loss: 0.0011616930132731795
step: 130, loss: 0.0019025264773517847
step: 140, loss: 0.009754570201039314
step: 150, loss: 0.0024347230792045593
step: 160, loss: 0.0012438296107575297
step: 170, loss: 0.001828474341891706
step: 180, loss: 0.0007162172696553171
step: 190, loss: 0.001496976474300027
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.7010869565217391, f1=0.7247956403269755,
best_f1=0.7374301675977653
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019773433450609446
step: 10, loss: 0.001206140615977347
step: 20, loss: 0.0026463570538908243
step: 30, loss: 0.051257483661174774
step: 40, loss: 0.0018766724970191717
step: 50, loss: 0.0009234769968315959
step: 60, loss: 0.0006796472007408738
step: 70, loss: 0.03241594880819321
step: 80, loss: 0.0007031364366412163
step: 90, loss: 0.0515257902443409
step: 100, loss: 0.0009793840581551194
step: 110, loss: 0.001675893785431981
step: 120, loss: 0.007974301464855671
step: 130, loss: 0.0010085513349622488
step: 140, loss: 0.00019486070959828794
step: 150, loss: 0.034983690828084946
step: 160, loss: 0.0013999751536175609
step: 170, loss: 0.19792276620864868
step: 180, loss: 0.00036728958366438746
step: 190, loss: 0.016618860885500908
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```



```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.7335243553008597, f1=0.7130434782608697,
best_f1=0.7374301675977653
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004274782841093838
step: 10, loss: 0.0004863630747422576
step: 20, loss: 0.00039096540422178805
step: 30, loss: 0.0013926810352131724
step: 40, loss: 0.020153552293777466
step: 50, loss: 0.0006904887268319726
step: 60, loss: 0.007154626771807671
step: 70, loss: 0.0019147576531395316
step: 80, loss: 0.0003523962805047631
step: 90, loss: 0.1475667953491211
step: 100, loss: 0.007745903916656971
step: 110, loss: 0.06712713837623596
step: 120, loss: 0.02405606582760811
step: 130, loss: 0.0006360786501318216
step: 140, loss: 0.0007869320106692612
step: 150, loss: 0.0002753592561930418
step: 160, loss: 0.0005693226703442633
step: 170, loss: 0.001609639497473836
step: 180, loss: 0.004014559555798769
step: 190, loss: 0.0015046967891976237
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.7272727272727272, f1=0.7411167512690356,
best_f1=0.7374301675977653
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00028401502640917897
step: 10, loss: 0.007403762545436621
step: 20, loss: 0.0018231066642329097
step: 30, loss: 0.00033864565193653107
step: 40, loss: 0.0009750601020641625
step: 50, loss: 0.000972036796156317
step: 60, loss: 0.001935016829520464
step: 70, loss: 0.0008226773352362216
step: 80, loss: 0.00020205907640047371
step: 90, loss: 0.001607624115422368
step: 100, loss: 0.00027051789220422506
step: 110, loss: 0.0003589804982766509
step: 120, loss: 0.009845581836998463

```

```

step: 130, loss: 0.02749299816787243
step: 140, loss: 0.0006995404255576432
step: 150, loss: 0.0005535880336537957
step: 160, loss: 0.0003133353893645108
step: 170, loss: 0.0015918859280645847
step: 180, loss: 0.005950368009507656
step: 190, loss: 0.0004947038250975311
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.7291666666666667, f1=0.7287234042553191,
best_f1=0.7374301675977653
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006621896289288998
step: 10, loss: 0.0015735016204416752
step: 20, loss: 0.0033553631510585546
step: 30, loss: 0.0006459542782977223
step: 40, loss: 0.0022623641416430473
step: 50, loss: 0.0014317313907667994
step: 60, loss: 0.000570896256249398
step: 70, loss: 0.009231018833816051
step: 80, loss: 0.0005633659893646836
step: 90, loss: 0.0005627265200018883
step: 100, loss: 0.0003750853065866977
step: 110, loss: 0.0004095544572919607
step: 120, loss: 0.0029530776664614677
step: 130, loss: 0.00026672970852814615
step: 140, loss: 0.0018973964033648372
step: 150, loss: 0.0003474641125649214
step: 160, loss: 0.009769369848072529
step: 170, loss: 0.00454504881054163
step: 180, loss: 0.0016243670834228396
step: 190, loss: 0.04161422327160835
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.7317073170731707, f1=0.732394366197183,
best_f1=0.7374301675977653
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005440758541226387
step: 10, loss: 0.0009578617755323648
step: 20, loss: 0.0008328538970090449
step: 30, loss: 0.09309446811676025
step: 40, loss: 0.00015415716916322708

```

```

step: 50, loss: 0.00022957853798288852
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.00010944147652480751
step: 70, loss: 0.00041418688488192856
step: 80, loss: 0.00020367966499179602
step: 90, loss: 0.0004936940968036652
step: 100, loss: 0.0004687706823460758
step: 110, loss: 0.000679333636071533
step: 120, loss: 0.01384652778506279
step: 130, loss: 0.011192303150892258
step: 140, loss: 0.0013767154887318611
step: 150, loss: 0.0007969338330440223
step: 160, loss: 0.004438660107553005
step: 170, loss: 0.00019935457385145128
step: 180, loss: 0.0015467065386474133
step: 190, loss: 0.001046139164827764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.7235142118863048, f1=0.7354497354497355,
best_f1=0.7374301675977653
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000592985947150737
step: 10, loss: 0.0024201960768550634
step: 20, loss: 0.0012222763616591692
step: 30, loss: 0.00025879524764604867
step: 40, loss: 0.0004083906242158264
step: 50, loss: 0.003567527746781707
step: 60, loss: 0.00019862016779370606
step: 70, loss: 0.00026331929257139564
step: 80, loss: 0.0001302661548834294
step: 90, loss: 0.0002895001962315291
step: 100, loss: 0.0012692040763795376
step: 110, loss: 0.0007928289123810828
step: 120, loss: 0.0004719884891528636
step: 130, loss: 0.00048557561240158975
step: 140, loss: 0.00038050534203648567
step: 150, loss: 0.00036275468301028013
step: 160, loss: 0.006889245007187128
step: 170, loss: 0.00014815706526860595
step: 180, loss: 0.004822852555662394
step: 190, loss: 0.0001075410473276861
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.7245657568238213, f1=0.7360406091370559,

```

```

best_f1=0.7374301675977653
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002494004089385271
step: 10, loss: 0.00041990558383986354
step: 20, loss: 0.0001592658954905346
step: 30, loss: 0.0002918969839811325
step: 40, loss: 0.0003750300966203213
step: 50, loss: 0.021119264885783195
step: 60, loss: 0.00017220959125552326
step: 70, loss: 0.00046101020416244864
step: 80, loss: 0.00027442839927971363
step: 90, loss: 0.0006992778507992625
step: 100, loss: 0.00021044013556092978
step: 110, loss: 0.0002105717285303399
step: 120, loss: 0.007444161456078291
step: 130, loss: 0.00033650730620138347
step: 140, loss: 0.0008245196077041328
step: 150, loss: 0.0012860188726335764
step: 160, loss: 0.003551019122824073
step: 170, loss: 0.00017648535140324384
step: 180, loss: 0.0011711451224982738
step: 190, loss: 0.014754247851669788
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.722772272727227, f1=0.7355163727959698,
best_f1=0.7374301675977653
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type       : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type       : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True

```

```

master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:09, 220.46it/s]
load_f1 = 0.7277486910994764
real_f1 = 0.7297297297297297
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:17, 253.17it/s]

```

## 2.2 BASELINE TEXTUAL

### 2.2.1 Abt-Buy - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
    --task Textual/Abt-Buy \
    --batch_size 32 \
    --max_len 256 \
    --lr 3e-5 \
    --n_epochs 15 \
    --finetuning \
    --lm distilbert \
    --fp16 \
    --save_model

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Abt-Buy.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
    --task Textual/Abt-Buy \
    --input_path input/input_small.jsonl \
    --output_path output/output_small.jsonl \

```

```
--lm distilbert \  
--use_gpu \  
--fp16 \  
--checkpoint_path checkpoints/
```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```
enabled          : True  
opt_level        : 02  
cast_model_type  : torch.float16  
patch_torch_functions : False  
keep_batchnorm_fp32 : True  
master_weights   : True  
loss_scale       : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```
enabled          : True  
opt_level        : 02  
cast_model_type  : torch.float16  
patch_torch_functions : False  
keep_batchnorm_fp32 : True  
master_weights   : True  
loss_scale       : dynamic
```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback.

Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0

/usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`.

In PyTorch 1.1.0 and later, you should call them in the opposite order:

`optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

"<https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>",  
UserWarning)

step: 0, loss: 0.8489307165145874

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

step: 10, loss: 0.22490908205509186

step: 20, loss: 0.15595832467079163

step: 30, loss: 0.23432329297065735

step: 40, loss: 0.3040045201778412

```

step: 50, loss: 0.3737771809101105
step: 60, loss: 0.4382357597351074
step: 70, loss: 0.305972695350647
step: 80, loss: 0.25423088669776917
step: 90, loss: 0.39144837856292725
step: 100, loss: 0.22318249940872192
step: 110, loss: 0.18734994530677795
step: 120, loss: 0.5498788356781006
step: 130, loss: 0.40470436215400696
step: 140, loss: 0.47772419452667236
step: 150, loss: 0.12645114958286285
step: 160, loss: 0.3308548927307129
step: 170, loss: 0.23047085106372833
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.4131944444444444, f1=0.38721136767317943,
best_f1=0.38721136767317943
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.38319769501686096
step: 10, loss: 0.1737217903137207
step: 20, loss: 0.30179744958877563
step: 30, loss: 0.29769545793533325
step: 40, loss: 0.17296546697616577
step: 50, loss: 0.22167228162288666
step: 60, loss: 0.13586516678333282
step: 70, loss: 0.25888076424598694
step: 80, loss: 0.09415074437856674
step: 90, loss: 0.10474631190299988
step: 100, loss: 0.15909627079963684
step: 110, loss: 0.1757412701845169
step: 120, loss: 0.06420145183801651
step: 130, loss: 0.0600828155875206
step: 140, loss: 0.0417439267039299
step: 150, loss: 0.12093179672956467
step: 160, loss: 0.13295328617095947
step: 170, loss: 0.12148317694664001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or

```

```

loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.7688888888888888, f1=0.7537473233404711,
best_f1=0.7537473233404711
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04561660438776016
step: 10, loss: 0.06639055907726288
step: 20, loss: 0.051056403666734695
step: 30, loss: 0.2815548777580261
step: 40, loss: 0.042075395584106445
step: 50, loss: 0.19024258852005005
step: 60, loss: 0.06395860761404037
step: 70, loss: 0.0600978322327137
step: 80, loss: 0.01680469885468483
step: 90, loss: 0.08199220895767212
step: 100, loss: 0.033337388187646866
step: 110, loss: 0.03607776761054993
step: 120, loss: 0.005265233106911182
step: 130, loss: 0.13774465024471283
step: 140, loss: 0.009562828578054905
step: 150, loss: 0.06433422863483429
step: 160, loss: 0.0392654649913311
step: 170, loss: 0.09364711493253708
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.7732696897374701, f1=0.8, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05370940640568733
step: 10, loss: 0.07002399861812592
step: 20, loss: 0.021722039207816124
step: 30, loss: 0.017373396083712578
step: 40, loss: 0.004317157901823521
step: 50, loss: 0.00664540147408843
step: 60, loss: 0.10350434482097626
step: 70, loss: 0.04580375924706459
step: 80, loss: 0.06860794126987457
step: 90, loss: 0.040220677852630615
step: 100, loss: 0.04306374862790108
step: 110, loss: 0.026790078729391098

```



```

step: 120, loss: 0.09780112653970718
step: 130, loss: 0.03336017578840256
step: 140, loss: 0.009348543360829353
step: 150, loss: 0.006026128772646189
step: 160, loss: 0.0692431852221489
step: 170, loss: 0.08481425791978836
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.7587822014051522, f1=0.7780320366132724, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04969645291566849
step: 10, loss: 0.07765452563762665
step: 20, loss: 0.021834278479218483
step: 30, loss: 0.02481267787516117
step: 40, loss: 0.05280400812625885
step: 50, loss: 0.008447381667792797
step: 60, loss: 0.00258445180952549
step: 70, loss: 0.13680043816566467
step: 80, loss: 0.006829500664025545
step: 90, loss: 0.05160653591156006
step: 100, loss: 0.018348684534430504
step: 110, loss: 0.0275360569357872
step: 120, loss: 0.07469374686479568
step: 130, loss: 0.006403895560652018
step: 140, loss: 0.02338656224310398
step: 150, loss: 0.011248331516981125
step: 160, loss: 0.0070044416934251785
step: 170, loss: 0.029249340295791626
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.7589498806682576, f1=0.7865707434052758, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008603558875620365
step: 10, loss: 0.05254874378442764
step: 20, loss: 0.0035698043648153543
step: 30, loss: 0.11907480657100677
step: 40, loss: 0.031271420419216156
step: 50, loss: 0.036825086921453476
step: 60, loss: 0.014673667028546333
step: 70, loss: 0.008721613325178623
step: 80, loss: 0.02039545215666294
step: 90, loss: 0.0032505160197615623

```

```

step: 100, loss: 0.13839247822761536
step: 110, loss: 0.0037459249142557383
step: 120, loss: 0.004433546215295792
step: 130, loss: 0.10337087512016296
step: 140, loss: 0.0033471332862973213
step: 150, loss: 0.12551164627075195
step: 160, loss: 0.09111866354942322
step: 170, loss: 0.01622362993657589
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.7666666666666666, f1=0.771764705882353, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026468231808394194
step: 10, loss: 0.002390149747952819
step: 20, loss: 0.062416478991508484
step: 30, loss: 0.004382003098726273
step: 40, loss: 0.013895680196583271
step: 50, loss: 0.0028650069143623114
step: 60, loss: 0.22209107875823975
step: 70, loss: 0.00744191138073802
step: 80, loss: 0.0022752166260033846
step: 90, loss: 0.014269627630710602
step: 100, loss: 0.002201530383899808
step: 110, loss: 0.0009450605139136314
step: 120, loss: 0.09682019054889679
step: 130, loss: 0.05415430665016174
step: 140, loss: 0.037665944546461105
step: 150, loss: 0.07540495693683624
step: 160, loss: 0.0021085471380501986
step: 170, loss: 0.22377130389213562
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.7634660421545668, f1=0.7414187643020594, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05407244712114334
step: 10, loss: 0.008485469967126846
step: 20, loss: 0.002239321591332555
step: 30, loss: 0.04710742458701134
step: 40, loss: 0.008771102875471115
step: 50, loss: 0.0020219217985868454
step: 60, loss: 0.01916099339723587
step: 70, loss: 0.016662878915667534

```

```

step: 80, loss: 0.002171084051951766
step: 90, loss: 0.02050732634961605
step: 100, loss: 0.010947974398732185
step: 110, loss: 0.021554116159677505
step: 120, loss: 0.0031748157925903797
step: 130, loss: 0.005663472227752209
step: 140, loss: 0.0016404450871050358
step: 150, loss: 0.07884878665208817
step: 160, loss: 0.00505125941708684
step: 170, loss: 0.0024230917915701866
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.764857881136951, f1=0.7841191066997518, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028356558177620173
step: 10, loss: 0.006010175682604313
step: 20, loss: 0.09136330336332321
step: 30, loss: 0.008661988191306591
step: 40, loss: 0.004533069208264351
step: 50, loss: 0.0059312572702765465
step: 60, loss: 0.0004679673002101481
step: 70, loss: 0.001544760656543076
step: 80, loss: 0.03312816098332405
step: 90, loss: 0.08531595021486282
step: 100, loss: 0.021061774343252182
step: 110, loss: 0.00407209200784564
step: 120, loss: 0.052038829773664474
step: 130, loss: 0.00041202432475984097
step: 140, loss: 0.005077302921563387
step: 150, loss: 0.009283571504056454
step: 160, loss: 0.033372342586517334
step: 170, loss: 0.002342880703508854
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.7706855791962176, f1=0.7906976744186047, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02248178981244564
step: 10, loss: 0.023593183606863022
step: 20, loss: 0.00044551154132932425
step: 30, loss: 0.015751739963889122
step: 40, loss: 0.0036901042331010103
step: 50, loss: 0.009757257997989655

```

```

step: 60, loss: 0.007700871676206589
step: 70, loss: 0.002987154293805361
step: 80, loss: 0.01408130768686533
step: 90, loss: 0.035185474902391434
step: 100, loss: 0.014063810929656029
step: 110, loss: 0.0020093433558940887
step: 120, loss: 0.0074689956381917
step: 130, loss: 0.008622339926660061
step: 140, loss: 0.001230994239449501
step: 150, loss: 0.0005484294961206615
step: 160, loss: 0.00013193253835197538
step: 170, loss: 0.004922097083181143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.7570093457943925, f1=0.7790432801822323, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021614364231936634
step: 10, loss: 0.022000566124916077
step: 20, loss: 0.07223468273878098
step: 30, loss: 0.002677659271284938
step: 40, loss: 0.0022121029905974865
step: 50, loss: 0.0075003039091825485
step: 60, loss: 0.00044112856267020106
step: 70, loss: 0.0002799195353873074
step: 80, loss: 0.0002824975526891649
step: 90, loss: 0.0010061439825221896
step: 100, loss: 0.0002974419039674103
step: 110, loss: 0.007935088127851486
step: 120, loss: 0.0887257307767868
step: 130, loss: 0.0013141883537173271
step: 140, loss: 0.04518783837556839
step: 150, loss: 0.015689712017774582
step: 160, loss: 0.0042230417020618916
step: 170, loss: 0.07694907486438751
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.7688679245283018, f1=0.7924528301886793, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009162419475615025
step: 10, loss: 0.019221624359488487
step: 20, loss: 0.05551890656352043
step: 30, loss: 0.0008832654566504061

```

```

step: 40, loss: 0.0012376991799101233
step: 50, loss: 0.00025831907987594604
step: 60, loss: 0.00024791507166810334
step: 70, loss: 0.004104292951524258
step: 80, loss: 0.053046274930238724
step: 90, loss: 0.009458135813474655
step: 100, loss: 0.00020932125335093588
step: 110, loss: 0.00044769461965188384
step: 120, loss: 0.0032681741286069155
step: 130, loss: 0.005127121694386005
step: 140, loss: 0.002270022640004754
step: 150, loss: 0.0006475072004832327
step: 160, loss: 0.010194845497608185
step: 170, loss: 0.0016474107978865504
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.7433155080213903, f1=0.7917737789203084, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012602756032720208
step: 10, loss: 0.06773052364587784
step: 20, loss: 0.0008241793257184327
step: 30, loss: 0.00018577827722765505
step: 40, loss: 0.0005717705935239792
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 50, loss: 0.005468354560434818
step: 60, loss: 0.003995097707957029
step: 70, loss: 0.02652771584689617
step: 80, loss: 0.0005569986533373594
step: 90, loss: 0.0002548234770074487
step: 100, loss: 0.0014455518685281277
step: 110, loss: 0.00021670793648809195
step: 120, loss: 0.0004896628088317811
step: 130, loss: 0.00031507143285125494
step: 140, loss: 0.0001676780666457489
step: 150, loss: 0.0006159176700748503
step: 160, loss: 0.0004165160935372114
step: 170, loss: 0.0005321121425367892
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.7469879518072288, f1=0.7905882352941176, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00020725594367831945

```

```

step: 10, loss: 0.00014850855222903192
step: 20, loss: 0.0007098073838278651
step: 30, loss: 0.00026232886011712253
step: 40, loss: 0.0003352465573698282
step: 50, loss: 0.007460605353116989
step: 60, loss: 0.0009636470349505544
step: 70, loss: 0.0018357463413849473
step: 80, loss: 0.0019375465344637632
step: 90, loss: 0.00019245874136686325
step: 100, loss: 0.00020918289374094456
step: 110, loss: 0.0011268740054219961
step: 120, loss: 0.050687436014413834
step: 130, loss: 0.010113094933331013
step: 140, loss: 0.011952316388487816
step: 150, loss: 0.0031200286466628313
step: 160, loss: 0.0481608547270298
step: 170, loss: 0.0003170035779476166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.7517730496453902, f1=0.7925407925407925, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004721107252407819
step: 10, loss: 0.0037888463120907545
step: 20, loss: 0.030511528253555298
step: 30, loss: 0.0066140396520495415
step: 40, loss: 0.002287638373672962
step: 50, loss: 0.0012086943024769425
step: 60, loss: 0.0005806998815387487
step: 70, loss: 0.0008880218374542892
step: 80, loss: 0.0003853060188703239
step: 90, loss: 0.00014269792882259935
step: 100, loss: 0.00014905663556419313
step: 110, loss: 0.00031474934075959027
step: 120, loss: 0.003972803242504597
step: 130, loss: 0.0002692951529752463
step: 140, loss: 0.00018578267190605402
step: 150, loss: 0.0335942804813385
step: 160, loss: 0.000132492117700167
step: 170, loss: 0.0004760223382618278
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.7461139896373058, f1=0.7969924812030076, best_f1=0.8
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32

```

master weights.

Defaults for this optimization level are:

```
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

1916it [00:04, 464.06it/s]

load\_f1 = 0.7696202531645568

real\_f1 = 0.7600000000000001

0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

4398it [00:11, 384.62it/s]

## 2.3 BASELINE DIRTY

### 2.3.1 DBLP-ACM - Running the matcher

```
[ ]: # Step 1:
      # train the wdc model with the save_model flag on
      # it will produce two files *_dev.pt and *_test.pt which are the
      # best checkpoints on the dev set and the test set.
      !CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
        --task Dirty/DBLP-ACM \
```

```

--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm distilbert \
--fp16 \
--save_model

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-ACM.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Dirty/DBLP-ACM \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm distilbert \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback.

Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning:

An input tensor was not cuda.



```

    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.8098803758621216
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4843759536743164
step: 20, loss: 0.6060479879379272
step: 30, loss: 0.4809229075908661
step: 40, loss: 0.3412077724933624
step: 50, loss: 0.0885297879576683
step: 60, loss: 0.17218032479286194
step: 70, loss: 0.16846922039985657
step: 80, loss: 0.2686627507209778
step: 90, loss: 0.013877779245376587
step: 100, loss: 0.09864412248134613
step: 110, loss: 0.0792241245508194
step: 120, loss: 0.024758653715252876
step: 130, loss: 0.016454054042696953
step: 140, loss: 0.04830991104245186
step: 150, loss: 0.15778377652168274
step: 160, loss: 0.18427498638629913
step: 170, loss: 0.010586201213300228
step: 180, loss: 0.021227670833468437
step: 190, loss: 0.02272905223071575
step: 200, loss: 0.056889358907938004
step: 210, loss: 0.039657674729824066
step: 220, loss: 0.023035649210214615
step: 230, loss: 0.01052690390497446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9501661129568106, f1=0.9551569506726457,
best_f1=0.9551569506726457
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.029018651694059372
step: 10, loss: 0.0011114564258605242
step: 20, loss: 0.005876426585018635
step: 30, loss: 0.014950858429074287
step: 40, loss: 0.0510282889008522
step: 50, loss: 0.038898009806871414
step: 60, loss: 0.01413451787084341
step: 70, loss: 0.01736215129494667
step: 80, loss: 0.02043703943490982
step: 90, loss: 0.007238356396555901
step: 100, loss: 0.1189713254570961
step: 110, loss: 0.1532433182001114
step: 120, loss: 0.013406489975750446
step: 130, loss: 0.002468266524374485
step: 140, loss: 0.25107213854789734
step: 150, loss: 0.09223975241184235
step: 160, loss: 0.01762576214969158
step: 170, loss: 0.07297798991203308
step: 180, loss: 0.010877995751798153
step: 190, loss: 0.15714633464813232
step: 200, loss: 0.039819248020648956
step: 210, loss: 0.04076077416539192
step: 220, loss: 0.0027576342690736055
step: 230, loss: 0.0064587402157485485
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9545957918050941, f1=0.9579646017699115,
best_f1=0.9579646017699115
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09659488499164581
step: 10, loss: 0.08015473186969757
step: 20, loss: 0.04196477681398392
step: 30, loss: 0.00882739294320345
step: 40, loss: 0.016048546880483627
step: 50, loss: 0.005937475711107254
step: 60, loss: 0.07017526775598526
step: 70, loss: 0.0021048325579613447
step: 80, loss: 0.0020548808388412
step: 90, loss: 0.06837659329175949
step: 100, loss: 0.008141462691128254
step: 110, loss: 0.03802843764424324

```

```

step: 120, loss: 0.04693177342414856
step: 130, loss: 0.006192310713231564
step: 140, loss: 0.006009404081851244
step: 150, loss: 0.017610061913728714
step: 160, loss: 0.017286287620663643
step: 170, loss: 0.0034355446696281433
step: 180, loss: 0.011471421457827091
step: 190, loss: 0.039282869547605515
step: 200, loss: 0.004190287087112665
step: 210, loss: 0.01759277656674385
step: 220, loss: 0.09326699376106262
step: 230, loss: 0.028195679187774658
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.9530201342281879, f1=0.9609810479375697,
best_f1=0.9579646017699115
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007204024586826563
step: 10, loss: 0.010121437720954418
step: 20, loss: 0.004083220846951008
step: 30, loss: 0.001501666847616434
step: 40, loss: 0.011639755219221115
step: 50, loss: 0.0026174280792474747
step: 60, loss: 0.0009735851781442761
step: 70, loss: 0.02687028795480728
step: 80, loss: 0.15254636108875275
step: 90, loss: 0.04413490742444992
step: 100, loss: 0.008988137356936932
step: 110, loss: 0.057072028517723083
step: 120, loss: 0.03139401227235794
step: 130, loss: 0.018710825592279434
step: 140, loss: 0.0005914228386245668
step: 150, loss: 0.01794315129518509
step: 160, loss: 0.009363760240375996
step: 170, loss: 0.0013237121747806668
step: 180, loss: 0.11635919660329819
step: 190, loss: 0.005928440950810909
step: 200, loss: 0.0019546381663531065
step: 210, loss: 0.01423337496817112
step: 220, loss: 0.01723167486488819
step: 230, loss: 0.001166021334938705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.9501661129568106, f1=0.954895489548955,

```

```

best_f1=0.9579646017699115
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002083214931190014
step: 10, loss: 0.001457000500522554
step: 20, loss: 0.0010554493637755513
step: 30, loss: 0.0004030624695587903
step: 40, loss: 0.0012620645575225353
step: 50, loss: 0.01066013053059578
step: 60, loss: 0.000605225854087621
step: 70, loss: 0.00038914912147447467
step: 80, loss: 0.0010359829757362604
step: 90, loss: 0.05490747094154358
step: 100, loss: 0.001772490912117064
step: 110, loss: 0.0005687451339326799
step: 120, loss: 0.029810409992933273
step: 130, loss: 0.1205463781952858
step: 140, loss: 0.0004375393909867853
step: 150, loss: 0.00017453636974096298
step: 160, loss: 0.003607360180467367
step: 170, loss: 0.05391814559698105
step: 180, loss: 0.0025366886984556913
step: 190, loss: 0.00046553683932870626
step: 200, loss: 0.0003775535151362419
step: 210, loss: 0.00026825195527635515
step: 220, loss: 0.01679699681699276
step: 230, loss: 0.00045946796308271587
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.9555555555555556, f1=0.966740576496674,
best_f1=0.966740576496674
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000993414781987667
step: 10, loss: 0.00032648612977936864
step: 20, loss: 0.0006563798524439335
step: 30, loss: 0.0002514658553991467
step: 40, loss: 0.07904700189828873
step: 50, loss: 0.03416551649570465
step: 60, loss: 0.1226535364985466
step: 70, loss: 0.00043617701157927513

```

```

step: 80, loss: 0.0023290691897273064
step: 90, loss: 0.0006153563153930008
step: 100, loss: 0.004518196918070316
step: 110, loss: 0.04012739285826683
step: 120, loss: 0.016812246292829514
step: 130, loss: 0.0047608851455152035
step: 140, loss: 0.005484471097588539
step: 150, loss: 0.024677900597453117
step: 160, loss: 0.05187331512570381
step: 170, loss: 0.0005109312478452921
step: 180, loss: 0.0010930189164355397
step: 190, loss: 0.2434828281402588
step: 200, loss: 0.0803365632891655
step: 210, loss: 0.010603142902255058
step: 220, loss: 0.0010207779705524445
step: 230, loss: 0.0016554604517295957
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 6: dev_f1=0.9576837416481068, f1=0.960352422907489,
best_f1=0.960352422907489
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16330474615097046
step: 10, loss: 0.006674927659332752
step: 20, loss: 0.00041051217704080045
step: 30, loss: 0.0012076995335519314
step: 40, loss: 0.001276715425774455
step: 50, loss: 0.004031925927847624
step: 60, loss: 0.10913451761007309
step: 70, loss: 0.027382155880331993
step: 80, loss: 0.00020342657808214426
step: 90, loss: 0.000852477562148124
step: 100, loss: 0.0006421747384592891
step: 110, loss: 0.0003097277949564159
step: 120, loss: 0.0035619325935840607
step: 130, loss: 0.0003791683411691338
step: 140, loss: 0.00016106704424601048
step: 150, loss: 0.0013778797583654523
step: 160, loss: 0.000629146583378315
step: 170, loss: 0.0002623272011987865
step: 180, loss: 0.00033507641637697816
step: 190, loss: 0.09545746445655823

```

```

step: 200, loss: 0.001516564516350627
step: 210, loss: 0.0018826471641659737
step: 220, loss: 0.01771785318851471
step: 230, loss: 0.03217030689120293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9535673839184597, f1=0.966740576496674,
best_f1=0.960352422907489
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014306876808404922
step: 10, loss: 0.013119221664965153
step: 20, loss: 0.00040745490696281195
step: 30, loss: 0.0004725004837382585
step: 40, loss: 0.0006042348686605692
step: 50, loss: 0.009194985032081604
step: 60, loss: 0.000874606950674206
step: 70, loss: 0.0007208282477222383
step: 80, loss: 0.0002948245091829449
step: 90, loss: 0.025309531018137932
step: 100, loss: 0.00020949241297785193
step: 110, loss: 0.0003032223030459136
step: 120, loss: 9.492416575085372e-05
step: 130, loss: 0.005085607059299946
step: 140, loss: 0.0004858847532887012
step: 150, loss: 0.00011628103675320745
step: 160, loss: 0.00044649263145402074
step: 170, loss: 0.003481200896203518
step: 180, loss: 0.00032070267479866743
step: 190, loss: 0.00019577618513721973
step: 200, loss: 0.0022810630034655333
step: 210, loss: 0.0006119449390098453
step: 220, loss: 0.005200034007430077
step: 230, loss: 0.14601504802703857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9551569506726457, f1=0.9645232815964524,
best_f1=0.960352422907489
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012577961198985577
step: 10, loss: 0.0015031477669253945
step: 20, loss: 0.0033248334657400846
step: 30, loss: 0.0004796420398633927

```

```

step: 40, loss: 0.0003091947583016008
step: 50, loss: 0.0036551132798194885
step: 60, loss: 0.00029744021594524384
step: 70, loss: 0.00027932279044762254
step: 80, loss: 0.01902041770517826
step: 90, loss: 0.0013744037132710218
step: 100, loss: 0.007294542621821165
step: 110, loss: 0.0002403288963250816
step: 120, loss: 0.054251305758953094
step: 130, loss: 0.0002484805590938777
step: 140, loss: 0.001247689826413989
step: 150, loss: 0.00011176755651831627
step: 160, loss: 0.0024453215301036835
step: 170, loss: 0.00024022658180911094
step: 180, loss: 0.0002937066601589322
step: 190, loss: 0.00012239243369549513
step: 200, loss: 0.00020722592307720333
step: 210, loss: 0.0001729616487864405
step: 220, loss: 0.015139463357627392
step: 230, loss: 0.0012866035103797913
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.9578713968957872, f1=0.9592061742006616,
best_f1=0.9592061742006616
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00020757464517373592
step: 10, loss: 0.00030159050947986543
step: 20, loss: 0.0011715888977050781
step: 30, loss: 0.0001945073454407975
step: 40, loss: 0.00020124319416936487
step: 50, loss: 0.0070341783575713634
step: 60, loss: 0.021565807983279228
step: 70, loss: 0.00027040179702453315
step: 80, loss: 0.0002585419570095837
step: 90, loss: 0.00011739708861568943
step: 100, loss: 0.0025229116436094046
step: 110, loss: 0.00021373546042013913
step: 120, loss: 0.012734748423099518
step: 130, loss: 0.0009366328013129532
step: 140, loss: 0.002061971463263035
step: 150, loss: 0.0001698611449683085

```

```

step: 160, loss: 0.000161708245286718
step: 170, loss: 0.00013470185513142496
step: 180, loss: 0.00022240597172640264
step: 190, loss: 0.00015974105917848647
step: 200, loss: 0.0008764532976783812
step: 210, loss: 0.00031688984017819166
step: 220, loss: 0.0005197296850383282
step: 230, loss: 0.0007141599780879915
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9535673839184597, f1=0.9654403567447045,
best_f1=0.9592061742006616
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00040792603977024555
step: 10, loss: 0.0022070377599447966
step: 20, loss: 0.00023531500482931733
step: 30, loss: 0.0029429136775434017
step: 40, loss: 0.0041256980039179325
step: 50, loss: 0.00508972629904747
step: 60, loss: 0.0002647695946507156
step: 70, loss: 0.00030129513470456004
step: 80, loss: 0.0010775706032291055
step: 90, loss: 0.00013698058319278061
step: 100, loss: 0.00037172946031205356
step: 110, loss: 0.0016106563853099942
step: 120, loss: 0.0007884716033004224
step: 130, loss: 0.00016865726502146572
step: 140, loss: 0.00017873937031254172
step: 150, loss: 0.00012987438822165132
step: 160, loss: 0.0047872536815702915
step: 170, loss: 0.019982900470495224
step: 180, loss: 0.00024386204313486814
step: 190, loss: 0.00022603363322559744
step: 200, loss: 0.00032407278195023537
step: 210, loss: 0.00013752351514995098
step: 220, loss: 0.00023677876743022352
step: 230, loss: 0.0050015319138765335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 11: dev_f1=0.9591836734693877, f1=0.9664429530201343,

```



```

best_f1=0.9664429530201343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008166201878339052
step: 10, loss: 0.002237082691863179
step: 20, loss: 0.00010042786743724719
step: 30, loss: 0.0011146770557388663
step: 40, loss: 0.00013165961718186736
step: 50, loss: 0.00021046536858193576
step: 60, loss: 0.008201532065868378
step: 70, loss: 0.0006756295333616436
step: 80, loss: 0.00010477657633600757
step: 90, loss: 0.0007245267042890191
step: 100, loss: 5.894191417610273e-05
step: 110, loss: 8.561574941268191e-05
step: 120, loss: 0.00011092852946603671
step: 130, loss: 0.00018853286746889353
step: 140, loss: 7.44529752410017e-05
step: 150, loss: 5.0865211960626766e-05
step: 160, loss: 0.008915103040635586
step: 170, loss: 0.00010532443411648273
step: 180, loss: 0.00011655729758786038
step: 190, loss: 9.663078526500612e-05
step: 200, loss: 0.00578945642337203
step: 210, loss: 0.00012441047874744982
step: 220, loss: 0.004219613503664732
step: 230, loss: 9.6033705631271e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9569160997732427, f1=0.9621380846325166,
best_f1=0.9664429530201343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.682493403786793e-05
step: 10, loss: 0.0009848843328654766
step: 20, loss: 0.00012706410780083388
step: 30, loss: 0.0029702361207455397
step: 40, loss: 0.00013775956176687032
step: 50, loss: 9.502852481091395e-05
step: 60, loss: 7.096474291756749e-05
step: 70, loss: 7.571239257231355e-05
step: 80, loss: 0.000231542915571481
step: 90, loss: 0.0001179927057819441
step: 100, loss: 8.350310963578522e-05
step: 110, loss: 8.364505629288033e-05

```

```

step: 120, loss: 0.00020338126341812313
step: 130, loss: 0.00011857831850647926
step: 140, loss: 0.0005133628146722913
step: 150, loss: 0.060934342443943024
step: 160, loss: 6.672075687674806e-05
step: 170, loss: 0.0005451897741295397
step: 180, loss: 0.00010610395111143589
step: 190, loss: 0.00027688569389283657
step: 200, loss: 8.796501060714945e-05
step: 210, loss: 0.00010075853060698137
step: 220, loss: 0.00015105603961274028
step: 230, loss: 6.449691136367619e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9591160220994475, f1=0.9613259668508287,
best_f1=0.9664429530201343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.832329331198707e-05
step: 10, loss: 3.955647480324842e-05
step: 20, loss: 0.00022215073113329709
step: 30, loss: 9.303921979153529e-05
step: 40, loss: 4.923688175040297e-05
step: 50, loss: 0.00023750416585244238
step: 60, loss: 5.4163574532140046e-05
step: 70, loss: 6.618844054173678e-05
step: 80, loss: 0.00013615125499200076
step: 90, loss: 0.00011982952855760232
step: 100, loss: 0.00043617942719720304
step: 110, loss: 0.0004200577095616609
step: 120, loss: 4.767209611600265e-05
step: 130, loss: 0.00010554787877481431
step: 140, loss: 0.0001640645059524104
step: 150, loss: 8.842087117955089e-05
step: 160, loss: 4.516996705206111e-05
step: 170, loss: 0.00033162598265334964
step: 180, loss: 0.00018449327035341412
step: 190, loss: 0.00035088328877463937
step: 200, loss: 5.3074316383572295e-05
step: 210, loss: 0.027666844427585602
step: 220, loss: 0.00012157384480815381
step: 230, loss: 3.0460665584541857e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9559322033898305, f1=0.9632925472747497,

```

```

best_f1=0.9664429530201343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.66208507027477e-05
step: 10, loss: 9.540330211166292e-05
step: 20, loss: 7.060425559757277e-05
step: 30, loss: 0.0005899129901081324
step: 40, loss: 8.552515646442771e-05
step: 50, loss: 9.365703590447083e-05
step: 60, loss: 5.493800563272089e-05
step: 70, loss: 6.988342647673562e-05
step: 80, loss: 0.005331822205334902
step: 90, loss: 4.289034768589772e-05
step: 100, loss: 8.994136442197487e-05
step: 110, loss: 6.896664126543328e-05
step: 120, loss: 0.00011242310574743897
step: 130, loss: 7.85423326306045e-05
step: 140, loss: 0.000207658507861197
step: 150, loss: 9.016213880386204e-05
step: 160, loss: 7.513439049944282e-05
step: 170, loss: 4.0901624743128195e-05
step: 180, loss: 7.374215056188405e-05
step: 190, loss: 0.00019762362353503704
step: 200, loss: 5.743053770856932e-05
step: 210, loss: 0.006399247795343399
step: 220, loss: 0.00021711741283070296
step: 230, loss: 0.0001213955765706487
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9560315670800451, f1=0.9632925472747497,
best_f1=0.9664429530201343
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True

```

```

opt_level                : 02
cast_model_type          : torch.float16
patch_torch_functions    : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale               : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2473it [00:07, 316.45it/s]
load_f1 = 0.9581920903954803
real_f1 = 0.9572072072072072
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:10, 400.19it/s]

```

### 2.3.2 DBLP-GoogleScholar - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Dirty/DBLP-GoogleScholar \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm distilbert \
  --fp16 \
  --save_model

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-GoogleScholar/.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Dirty/DBLP-GoogleScholar \

```

```

--input_path input/input_small.jsonl \
--output_path output/output_small.jsonl \
--lm distilbert \
--use_gpu \
--fp16 \
--checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0 /usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`.

In PyTorch 1.1.0 and later, you should call them in the opposite order:

`optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

"<https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>", UserWarning)

step: 0, loss: 0.7941266298294067

step: 10, loss: 0.4010302722454071

step: 20, loss: 0.4844156801700592

step: 30, loss: 0.4486212134361267

step: 40, loss: 0.3998180627822876  
step: 50, loss: 0.23281477391719818  
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0  
step: 60, loss: 0.24504835903644562  
step: 70, loss: 0.20775850117206573  
step: 80, loss: 0.09816119819879532  
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0  
step: 90, loss: 0.20157063007354736  
step: 100, loss: 0.28308841586112976  
step: 110, loss: 0.09336906671524048  
step: 120, loss: 0.09273708611726761  
step: 130, loss: 0.029415138065814972  
step: 140, loss: 0.12553630769252777  
step: 150, loss: 0.048361290246248245  
step: 160, loss: 0.2594403028488159  
step: 170, loss: 0.18639220297336578  
step: 180, loss: 0.12529690563678741  
step: 190, loss: 0.04653630405664444  
step: 200, loss: 0.11306719481945038  
step: 210, loss: 0.09370004385709763  
step: 220, loss: 0.17781056463718414  
step: 230, loss: 0.17641140520572662  
step: 240, loss: 0.12560726702213287  
step: 250, loss: 0.07789665460586548  
step: 260, loss: 0.046323906630277634  
step: 270, loss: 0.01432714145630598  
step: 280, loss: 0.11552801728248596  
step: 290, loss: 0.14088454842567444  
step: 300, loss: 0.09438049793243408  
step: 310, loss: 0.1337401568889618  
step: 320, loss: 0.07213776558637619  
step: 330, loss: 0.12634123861789703  
step: 340, loss: 0.25908252596855164  
step: 350, loss: 0.09189695119857788  
step: 360, loss: 0.10829265415668488  
step: 370, loss: 0.1752208024263382  
step: 380, loss: 0.24722270667552948  
step: 390, loss: 0.10454975068569183  
step: 400, loss: 0.06172361597418785  
step: 410, loss: 0.015481388196349144  
step: 420, loss: 0.02084275707602501  
step: 430, loss: 0.09795690327882767  
step: 440, loss: 0.11803007870912552  
step: 450, loss: 0.13382774591445923  
step: 460, loss: 0.06529306620359421  
step: 470, loss: 0.23495861887931824  
step: 480, loss: 0.30583158135414124  
step: 490, loss: 0.0471004955470562

```

step: 500, loss: 0.04798964038491249
step: 510, loss: 0.1303468644618988
step: 520, loss: 0.07338813692331314
step: 530, loss: 0.05773342400789261
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9137291280148423, f1=0.9121184088806661,
best_f1=0.9121184088806661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13343067467212677
step: 10, loss: 0.17901049554347992
step: 20, loss: 0.08398057520389557
step: 30, loss: 0.02129601687192917
step: 40, loss: 0.007737387903034687
step: 50, loss: 0.04374343156814575
step: 60, loss: 0.09386318922042847
step: 70, loss: 0.18888357281684875
step: 80, loss: 0.012159080244600773
step: 90, loss: 0.02114175818860531
step: 100, loss: 0.3567405343055725
step: 110, loss: 0.028035709634423256
step: 120, loss: 0.07363545149564743
step: 130, loss: 0.021900292485952377
step: 140, loss: 0.01435847394168377
step: 150, loss: 0.03069121576845646
step: 160, loss: 0.01716659963130951
step: 170, loss: 0.14813768863677979
step: 180, loss: 0.005351528525352478
step: 190, loss: 0.047761086374521255
step: 200, loss: 0.061037782579660416
step: 210, loss: 0.06338929384946823
step: 220, loss: 0.16395612061023712
step: 230, loss: 0.09294719249010086
step: 240, loss: 0.12503932416439056
step: 250, loss: 0.0387783907353878
step: 260, loss: 0.023063138127326965
step: 270, loss: 0.6136698722839355
step: 280, loss: 0.38185614347457886
step: 290, loss: 0.10577530413866043
step: 300, loss: 0.06334920972585678
step: 310, loss: 0.027975600212812424

```

```

step: 320, loss: 0.1048547774553299
step: 330, loss: 0.030561575666069984
step: 340, loss: 0.01219206117093563
step: 350, loss: 0.11181531846523285
step: 360, loss: 0.050950150936841965
step: 370, loss: 0.010560438968241215
step: 380, loss: 0.11382189393043518
step: 390, loss: 0.018591850996017456
step: 400, loss: 0.11906645447015762
step: 410, loss: 0.0011392151936888695
step: 420, loss: 0.05969151481986046
step: 430, loss: 0.06033138558268547
step: 440, loss: 0.011307147331535816
step: 450, loss: 0.03035276010632515
step: 460, loss: 0.3923977315425873
step: 470, loss: 0.043049100786447525
step: 480, loss: 0.18338307738304138
step: 490, loss: 0.017452187836170197
step: 500, loss: 0.009326367639005184
step: 510, loss: 0.12701916694641113
step: 520, loss: 0.05746668204665184
step: 530, loss: 0.17164918780326843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9216504404265183, f1=0.9219330855018587,
best_f1=0.9219330855018587
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013770224526524544
step: 10, loss: 0.038760680705308914
step: 20, loss: 0.09986414760351181
step: 30, loss: 0.16572119295597076
step: 40, loss: 0.0029226012993603945
step: 50, loss: 0.05114312842488289
step: 60, loss: 0.031414810568094254
step: 70, loss: 0.011224630288779736
step: 80, loss: 0.16740526258945465
step: 90, loss: 0.08220197260379791
step: 100, loss: 0.007657116744667292
step: 110, loss: 0.08756505697965622
step: 120, loss: 0.007829218171536922
step: 130, loss: 0.012426642701029778

```



```

step: 140, loss: 0.012600177899003029
step: 150, loss: 0.006751951761543751
step: 160, loss: 0.02016112394630909
step: 170, loss: 0.003216997953131795
step: 180, loss: 0.03227423131465912
step: 190, loss: 0.013268960639834404
step: 200, loss: 0.0029644339811056852
step: 210, loss: 0.062074579298496246
step: 220, loss: 0.08561315387487411
step: 230, loss: 0.024047404527664185
step: 240, loss: 0.009979646652936935
step: 250, loss: 0.051714103668928146
step: 260, loss: 0.0030474690720438957
step: 270, loss: 0.0034184749238193035
step: 280, loss: 0.015647966414690018
step: 290, loss: 0.09825053811073303
step: 300, loss: 0.09107176959514618
step: 310, loss: 0.13847748935222626
step: 320, loss: 0.0777241587638855
step: 330, loss: 0.011106721125543118
step: 340, loss: 0.0011581730796024203
step: 350, loss: 0.023485733196139336
step: 360, loss: 0.04490990564227104
step: 370, loss: 0.003947594203054905
step: 380, loss: 0.043558694422245026
step: 390, loss: 0.01523859053850174
step: 400, loss: 0.047847554087638855
step: 410, loss: 0.056852787733078
step: 420, loss: 0.005411745049059391
step: 430, loss: 0.027295920997858047
step: 440, loss: 0.02871992625296116
step: 450, loss: 0.1543034166097641
step: 460, loss: 0.14197731018066406
step: 470, loss: 0.027519945055246353
step: 480, loss: 0.0036789746955037117
step: 490, loss: 0.03709834814071655
step: 500, loss: 0.025367802008986473
step: 510, loss: 0.0058129942044615746
step: 520, loss: 0.005794004071503878
step: 530, loss: 0.05185551941394806
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.9212669683257918, f1=0.924476797088262,
best_f1=0.9219330855018587
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 0.005204431712627411  
step: 10, loss: 0.017613736912608147  
step: 20, loss: 0.015173503197729588  
step: 30, loss: 0.02344265580177307  
step: 40, loss: 0.04572049528360367  
step: 50, loss: 0.02438226528465748  
step: 60, loss: 0.0015451916260644794  
step: 70, loss: 0.004580813925713301  
step: 80, loss: 0.08344369381666183  
step: 90, loss: 0.01656319387257099  
step: 100, loss: 0.04608011618256569  
step: 110, loss: 0.0046028089709579945  
step: 120, loss: 0.005308700725436211  
step: 130, loss: 0.06431976705789566  
step: 140, loss: 0.03780569136142731  
step: 150, loss: 0.008305344730615616  
step: 160, loss: 0.011315061710774899  
step: 170, loss: 0.0018855802481994033  
step: 180, loss: 0.005051225423812866  
step: 190, loss: 0.010723862797021866  
step: 200, loss: 0.00320791220292449  
step: 210, loss: 0.03351496160030365  
step: 220, loss: 0.005311222281306982  
step: 230, loss: 0.23501160740852356  
step: 240, loss: 0.043928056955337524  
step: 250, loss: 0.0008370814612135291  
step: 260, loss: 0.12217769026756287  
step: 270, loss: 0.021416887640953064  
step: 280, loss: 0.0008083348511718214  
step: 290, loss: 0.24109481275081635  
step: 300, loss: 0.0026326149236410856  
step: 310, loss: 0.0770338699221611  
step: 320, loss: 0.08467867970466614  
step: 330, loss: 0.05185973644256592  
step: 340, loss: 0.02773173525929451  
step: 350, loss: 0.05087165907025337  
step: 360, loss: 0.014230906032025814  
step: 370, loss: 0.016127554699778557  
step: 380, loss: 0.0009071369422599673  
step: 390, loss: 0.026672707870602608  
step: 400, loss: 0.013493170961737633  
step: 410, loss: 0.0011873060138896108  
step: 420, loss: 0.002261184388771653  
step: 430, loss: 0.014312149956822395  
step: 440, loss: 0.04158569127321243  
step: 450, loss: 0.004128775559365749  
step: 460, loss: 0.0009118773159570992  
step: 470, loss: 0.0017628230853006244

```

step: 480, loss: 0.0032405259553343058
step: 490, loss: 0.055761370807886124
step: 500, loss: 0.00807256530970335
step: 510, loss: 0.014028884470462799
step: 520, loss: 0.01210497971624136
step: 530, loss: 0.003358290297910571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.9207403891789273, f1=0.9097888675623801,
best_f1=0.9219330855018587
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001643202151171863
step: 10, loss: 0.010319647379219532
step: 20, loss: 0.007341546937823296
step: 30, loss: 0.001001891097985208
step: 40, loss: 0.07979676127433777
step: 50, loss: 0.0025179460644721985
step: 60, loss: 0.0030709945131093264
step: 70, loss: 0.001510357717052102
step: 80, loss: 0.0015929731307551265
step: 90, loss: 0.005270324181765318
step: 100, loss: 0.002592549892142415
step: 110, loss: 0.020201461389660835
step: 120, loss: 0.004480468574911356
step: 130, loss: 0.0018819186370819807
step: 140, loss: 0.00452408567070961
step: 150, loss: 0.0031500165350735188
step: 160, loss: 0.011018272489309311
step: 170, loss: 0.001090764650143683
step: 180, loss: 0.014665188267827034
step: 190, loss: 0.001262935227714479
step: 200, loss: 0.00494389096274972
step: 210, loss: 0.044330500066280365
step: 220, loss: 0.00277229817584157
step: 230, loss: 0.0008598034037277102
step: 240, loss: 0.018657490611076355
step: 250, loss: 0.0010276914108544588
step: 260, loss: 0.02211439609527588
step: 270, loss: 0.005503401160240173
step: 280, loss: 0.06489759683609009
step: 290, loss: 0.0006141445483081043
step: 300, loss: 0.012098919600248337
step: 310, loss: 0.00591938616707921
step: 320, loss: 0.0018592295236885548
step: 330, loss: 0.0008324990631081164

```

```

step: 340, loss: 0.004842331167310476
step: 350, loss: 0.005966451484709978
step: 360, loss: 0.00595754524692893
step: 370, loss: 0.0018647246761247516
step: 380, loss: 0.05855705961585045
step: 390, loss: 0.005195847246795893
step: 400, loss: 0.012626615352928638
step: 410, loss: 0.0006122299819253385
step: 420, loss: 0.005157703533768654
step: 430, loss: 0.0009545476641505957
step: 440, loss: 0.024693652987480164
step: 450, loss: 0.008665685541927814
step: 460, loss: 0.00589038897305727
step: 470, loss: 0.002591843483969569
step: 480, loss: 0.03259194269776344
step: 490, loss: 0.041814886033535004
step: 500, loss: 0.018895167857408524
step: 510, loss: 0.0038473871536552906
step: 520, loss: 0.005587236024439335
step: 530, loss: 0.010982371866703033
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.9236298292902065, f1=0.9118843199277, best_f1=0.9118843199277
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018694020109251142
step: 10, loss: 0.043508872389793396
step: 20, loss: 0.00042203409248031676
step: 30, loss: 0.01911010965704918
step: 40, loss: 0.002459883224219084
step: 50, loss: 0.02758009359240532
step: 60, loss: 0.00047240997082553804
step: 70, loss: 0.005447451025247574
step: 80, loss: 0.0015779477544128895
step: 90, loss: 0.004186585079878569
step: 100, loss: 0.002422966295853257
step: 110, loss: 0.0028867898508906364
step: 120, loss: 0.0017435471527278423
step: 130, loss: 0.0002684979699552059
step: 140, loss: 0.018894551321864128
step: 150, loss: 0.007660091854631901
step: 160, loss: 0.0002523054135963321

```

```

step: 170, loss: 0.011626854538917542
step: 180, loss: 0.00020460857194848359
step: 190, loss: 0.010611473582684994
step: 200, loss: 0.00019126168626826257
step: 210, loss: 0.00015139984316192567
step: 220, loss: 0.00011110257764812559
step: 230, loss: 0.0002256871957797557
step: 240, loss: 0.0010307555785402656
step: 250, loss: 0.0006390821072272956
step: 260, loss: 0.005980621092021465
step: 270, loss: 0.0005658636800944805
step: 280, loss: 0.003677238943055272
step: 290, loss: 0.0006678919307887554
step: 300, loss: 0.006988252513110638
step: 310, loss: 0.0005467177252285182
step: 320, loss: 0.0036496473476290703
step: 330, loss: 0.002503069117665291
step: 340, loss: 0.0007260905113071203
step: 350, loss: 0.05590752884745598
step: 360, loss: 0.009548911824822426
step: 370, loss: 0.002549865748733282
step: 380, loss: 0.0006271462771110237
step: 390, loss: 0.002033105120062828
step: 400, loss: 0.0001928978890646249
step: 410, loss: 0.0004093909519724548
step: 420, loss: 0.03839411213994026
step: 430, loss: 0.00046388644841499627
step: 440, loss: 0.001317776506766677
step: 450, loss: 0.0003269968437962234
step: 460, loss: 0.031570661813020706
step: 470, loss: 0.00036230345722287893
step: 480, loss: 0.005629044491797686
step: 490, loss: 0.00047820742474868894
step: 500, loss: 0.012722862884402275
step: 510, loss: 0.0002967157051898539
step: 520, loss: 0.0018856681417673826
step: 530, loss: 0.0033501298166811466
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9160871580899397, f1=0.9147940074906367,
best_f1=0.9118843199277
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.027187073603272438
step: 10, loss: 0.002855498343706131
step: 20, loss: 0.014451082795858383

```

step: 30, loss: 0.08971710503101349  
step: 40, loss: 0.014601105824112892  
step: 50, loss: 0.037453263998031616  
step: 60, loss: 0.0009041354642249644  
step: 70, loss: 0.0023108888417482376  
step: 80, loss: 0.00029117189114913344  
step: 90, loss: 0.0014396277256309986  
step: 100, loss: 0.0005045818979851902  
step: 110, loss: 0.015027357265353203  
step: 120, loss: 0.0004087587876711041  
step: 130, loss: 0.0010666828602552414  
step: 140, loss: 0.0022428384982049465  
step: 150, loss: 0.00012846163008362055  
step: 160, loss: 0.00014148573973216116  
step: 170, loss: 0.02180585451424122  
step: 180, loss: 0.00032912922324612737  
step: 190, loss: 0.00011443333642091602  
step: 200, loss: 0.000165397665114142  
step: 210, loss: 0.0025163067039102316  
step: 220, loss: 0.0038848239928483963  
step: 230, loss: 0.003986045252531767  
step: 240, loss: 0.004211201798170805  
step: 250, loss: 0.01602872833609581  
step: 260, loss: 0.0254294965416193  
step: 270, loss: 0.00021360287792049348  
step: 280, loss: 0.0718216598033905  
step: 290, loss: 0.031031813472509384  
step: 300, loss: 0.0017593428492546082  
step: 310, loss: 0.0004896640894003212  
step: 320, loss: 0.02354997582733631  
step: 330, loss: 0.00011730011465260759  
step: 340, loss: 0.07379680126905441  
step: 350, loss: 0.002004058100283146  
step: 360, loss: 0.0022331972140818834  
step: 370, loss: 0.0484110489487648  
step: 380, loss: 0.0015435429522767663  
step: 390, loss: 0.00037433920078910887  
step: 400, loss: 0.006131244823336601  
step: 410, loss: 0.01665003038942814  
step: 420, loss: 0.003651856677606702  
step: 430, loss: 0.007543259300291538  
step: 440, loss: 5.9764035540865734e-05  
step: 450, loss: 0.0001986611750908196  
step: 460, loss: 0.0003109889221377671  
step: 470, loss: 0.00019608053844422102  
step: 480, loss: 0.0001670478959567845  
step: 490, loss: 0.01982424594461918  
step: 500, loss: 7.179982640082017e-05

```

step: 510, loss: 0.000504210707731545
step: 520, loss: 0.13459137082099915
step: 530, loss: 0.00026907690335065126
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.9248716752216519, f1=0.9161592505854801,
best_f1=0.9161592505854801
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011339231859892607
step: 10, loss: 0.0006944983033463359
step: 20, loss: 0.037851061671972275
step: 30, loss: 0.0008969451300799847
step: 40, loss: 0.0007281905855052173
step: 50, loss: 0.00203043594956398
step: 60, loss: 0.00013355685223359615
step: 70, loss: 0.0002476298832334578
step: 80, loss: 0.00010399902384961024
step: 90, loss: 0.0003771260962821543
step: 100, loss: 0.001853506313636899
step: 110, loss: 0.00020476784266065806
step: 120, loss: 0.0185652244836092
step: 130, loss: 0.0010935120517387986
step: 140, loss: 4.300692671677098e-05
step: 150, loss: 0.019766390323638916
step: 160, loss: 0.00048304087249562144
step: 170, loss: 0.00015605735825374722
step: 180, loss: 6.990028487052768e-05
step: 190, loss: 0.0002912769850809127
step: 200, loss: 0.0002700735640246421
step: 210, loss: 0.09813268482685089
step: 220, loss: 0.00028661135002039373
step: 230, loss: 0.0007878720061853528
step: 240, loss: 0.00012303382391110063
step: 250, loss: 0.0005589311476796865
step: 260, loss: 0.007034163922071457
step: 270, loss: 5.620490628643893e-05
step: 280, loss: 0.0001255699316971004
step: 290, loss: 0.0003875017282553017
step: 300, loss: 0.00012665739632211626
step: 310, loss: 0.015520131215453148
step: 320, loss: 0.004350247327238321

```

```

step: 330, loss: 0.011240866035223007
step: 340, loss: 0.001316552865318954
step: 350, loss: 0.01925252191722393
step: 360, loss: 0.0001472742878831923
step: 370, loss: 0.0013021132908761501
step: 380, loss: 0.0008178098360076547
step: 390, loss: 0.0002679359749890864
step: 400, loss: 0.22027738392353058
step: 410, loss: 0.00830848328769207
step: 420, loss: 8.27887561172247e-05
step: 430, loss: 0.003138270927593112
step: 440, loss: 0.00041090563172474504
step: 450, loss: 0.0014814622700214386
step: 460, loss: 0.004047510679811239
step: 470, loss: 0.00011922147677978501
step: 480, loss: 0.027595164254307747
step: 490, loss: 0.0003854656533803791
step: 500, loss: 0.00038547193980775774
step: 510, loss: 0.00014789991837460548
step: 520, loss: 0.0006608807598240674
step: 530, loss: 0.00012213924492243677
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 8: dev_f1=0.9249658935879944, f1=0.9202566452795601,
best_f1=0.9202566452795601
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012135523866163567
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.00015421897114720196
step: 20, loss: 0.00022926070960238576
step: 30, loss: 0.01002549845725298
step: 40, loss: 0.0010820060269907117
step: 50, loss: 0.00023256278655026108
step: 60, loss: 0.00028950636624358594
step: 70, loss: 0.09046792984008789
step: 80, loss: 0.000522601418197155
step: 90, loss: 0.02220277488231659
step: 100, loss: 0.002272728132084012
step: 110, loss: 0.02437025122344494
step: 120, loss: 0.0001241370482603088
step: 130, loss: 0.0011550537310540676

```



```

step: 140, loss: 0.01394871436059475
step: 150, loss: 0.00035781311453320086
step: 160, loss: 0.0008731652051210403
step: 170, loss: 8.67454000399448e-05
step: 180, loss: 0.00011338093463564292
step: 190, loss: 0.06088130176067352
step: 200, loss: 0.00014924300194252282
step: 210, loss: 7.237218233058229e-05
step: 220, loss: 8.428994624409825e-05
step: 230, loss: 5.3578129154630005e-05
step: 240, loss: 0.003873933106660843
step: 250, loss: 0.00025896422448568046
step: 260, loss: 0.00014708380331285298
step: 270, loss: 0.00018459634156897664
step: 280, loss: 6.505749479401857e-05
step: 290, loss: 9.318684169556946e-05
step: 300, loss: 0.00010361838212702423
step: 310, loss: 0.0002499186375644058
step: 320, loss: 0.0007495586178265512
step: 330, loss: 7.097788329701871e-05
step: 340, loss: 0.041439179331064224
step: 350, loss: 0.0003144616784993559
step: 360, loss: 0.002552956808358431
step: 370, loss: 0.0004388547968119383
step: 380, loss: 7.223587454063818e-05
step: 390, loss: 0.00013419138849712908
step: 400, loss: 0.00046218070201575756
step: 410, loss: 0.00025039276806637645
step: 420, loss: 0.0001933535822900012
step: 430, loss: 0.0001850442058639601
step: 440, loss: 9.085487545235083e-05
step: 450, loss: 0.00013140629744157195
step: 460, loss: 0.001128419884480536
step: 470, loss: 0.0026292563416063786
step: 480, loss: 9.107038204092532e-05
step: 490, loss: 0.00015090632950887084
step: 500, loss: 0.0003117762680631131
step: 510, loss: 7.116035703802481e-05
step: 520, loss: 0.0009104013442993164
step: 530, loss: 0.00016221919213421643
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9068982151471296, f1=0.8937198067632851,
best_f1=0.9202566452795601
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 0.10635776072740555  
step: 10, loss: 5.701297413907014e-05  
step: 20, loss: 0.0215689018368721  
step: 30, loss: 0.0014928809832781553  
step: 40, loss: 0.00046050717355683446  
step: 50, loss: 7.538485806435347e-05  
step: 60, loss: 0.026126982644200325  
step: 70, loss: 0.04108475148677826  
step: 80, loss: 0.0019734082743525505  
step: 90, loss: 0.0024154859129339457  
step: 100, loss: 0.0008018465014174581  
step: 110, loss: 0.00034877884900197387  
step: 120, loss: 0.01411769725382328  
step: 130, loss: 0.004146266262978315  
step: 140, loss: 0.0005120875430293381  
step: 150, loss: 0.0012584890937432647  
step: 160, loss: 0.00039567076601088047  
step: 170, loss: 0.0002445615828037262  
step: 180, loss: 0.002211823593825102  
step: 190, loss: 0.0009067541104741395  
step: 200, loss: 0.0003121667541563511  
step: 210, loss: 0.00013452938583213836  
step: 220, loss: 0.003911548759788275  
step: 230, loss: 9.501448221271858e-05  
step: 240, loss: 8.095675002550706e-05  
step: 250, loss: 4.530032310867682e-05  
step: 260, loss: 0.0019016602309420705  
step: 270, loss: 0.0002558648120611906  
step: 280, loss: 0.00019254193466622382  
step: 290, loss: 0.00015685254766140133  
step: 300, loss: 0.2141934186220169  
step: 310, loss: 0.005449999123811722  
step: 320, loss: 0.000753294094465673  
step: 330, loss: 0.0004521292576100677  
step: 340, loss: 5.3753330576000735e-05  
step: 350, loss: 5.8598296163836494e-05  
step: 360, loss: 0.002452555578202009  
step: 370, loss: 0.0015881245490163565  
step: 380, loss: 7.672205538256094e-05  
step: 390, loss: 5.596154369413853e-05  
step: 400, loss: 8.100015111267567e-05  
step: 410, loss: 0.000624372623860836  
step: 420, loss: 0.00043300073593854904  
step: 430, loss: 0.00011008782894350588  
step: 440, loss: 0.0001435339218005538  
step: 450, loss: 0.008266893215477467  
step: 460, loss: 0.0026749013923108578  
step: 470, loss: 4.4812328269472346e-05

```

step: 480, loss: 4.900352723780088e-05
step: 490, loss: 0.015236110426485538
step: 500, loss: 0.0010816464200615883
step: 510, loss: 0.0002485472650732845
step: 520, loss: 4.550226731225848e-05
step: 530, loss: 6.15775934420526e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.918560606060606, f1=0.9047619047619048,
best_f1=0.9202566452795601
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.672768308315426e-05
step: 10, loss: 0.004535673651844263
step: 20, loss: 0.00014932926569599658
step: 30, loss: 0.0002209908707300201
step: 40, loss: 0.0001302729797316715
step: 50, loss: 8.983506995718926e-05
step: 60, loss: 0.00036750821163877845
step: 70, loss: 0.0011607403866946697
step: 80, loss: 0.0009522623149678111
step: 90, loss: 9.010593203129247e-05
step: 100, loss: 0.00012319118832238019
step: 110, loss: 8.612414239905775e-05
step: 120, loss: 0.0007494979654438794
step: 130, loss: 0.0009829988703131676
step: 140, loss: 7.995781197678298e-05
step: 150, loss: 0.00038547447184100747
step: 160, loss: 0.0002482068375684321
step: 170, loss: 0.00039781638770364225
step: 180, loss: 0.0001535962801426649
step: 190, loss: 9.359842806588858e-05
step: 200, loss: 0.0037824108731001616
step: 210, loss: 8.259514288511127e-05
step: 220, loss: 0.0001535944757051766
step: 230, loss: 0.0006001046276651323
step: 240, loss: 9.699423389974982e-05
step: 250, loss: 0.0011816928163170815
step: 260, loss: 0.00033608140074647963
step: 270, loss: 0.0017190460348501801
step: 280, loss: 6.145851511973888e-05
step: 290, loss: 0.008363917469978333
step: 300, loss: 0.012631671503186226
step: 310, loss: 0.00011471812467789277
step: 320, loss: 0.0015390977496281266
step: 330, loss: 0.00034024688648059964

```

```

step: 340, loss: 4.727484338218346e-05
step: 350, loss: 0.02152070589363575
step: 360, loss: 0.00019982349476777017
step: 370, loss: 2.7849549951497465e-05
step: 380, loss: 3.743623165064491e-05
step: 390, loss: 0.008236650377511978
step: 400, loss: 3.3090364013332874e-05
step: 410, loss: 0.0010986134875565767
step: 420, loss: 0.004901170264929533
step: 430, loss: 4.427876046975143e-05
step: 440, loss: 6.65579063934274e-05
step: 450, loss: 4.458955299924128e-05
step: 460, loss: 0.0008225191268138587
step: 470, loss: 0.017079869285225868
step: 480, loss: 0.0004138367366977036
step: 490, loss: 0.0007126957643777132
step: 500, loss: 7.455459126504138e-05
step: 510, loss: 8.042824629228562e-05
step: 520, loss: 5.171875818632543e-05
step: 530, loss: 3.817543256445788e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.922498825739784, f1=0.9113207547169812,
best_f1=0.9202566452795601
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016344841569662094
step: 10, loss: 6.833514635218307e-05
step: 20, loss: 6.774850044166669e-05
step: 30, loss: 5.630117084365338e-05
step: 40, loss: 0.00013580475933849812
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.05702956020832062
step: 60, loss: 0.00021423716680146754
step: 70, loss: 0.005855994299054146
step: 80, loss: 0.0019685165025293827
step: 90, loss: 9.190801210934296e-05
step: 100, loss: 0.00011127606558147818
step: 110, loss: 5.143110320204869e-05
step: 120, loss: 0.003451702417805791
step: 130, loss: 0.0004112066235393286
step: 140, loss: 5.9450234402902424e-05
step: 150, loss: 0.012126060202717781
step: 160, loss: 0.0002837203210219741
step: 170, loss: 0.00011342296056682244
step: 180, loss: 6.895478145452216e-05

```

```

step: 190, loss: 0.00010100136569235474
step: 200, loss: 8.95072880666703e-05
step: 210, loss: 0.00015951394743751734
step: 220, loss: 3.365334850968793e-05
step: 230, loss: 0.00046600797213613987
step: 240, loss: 0.00010910384298767895
step: 250, loss: 0.00078642024891451
step: 260, loss: 0.00018755717610474676
step: 270, loss: 7.945946708787233e-05
step: 280, loss: 6.234658940229565e-05
step: 290, loss: 0.0008483701967634261
step: 300, loss: 0.0004209227627143264
step: 310, loss: 7.768202340230346e-05
step: 320, loss: 0.00012576088192872703
step: 330, loss: 3.7087203963892534e-05
step: 340, loss: 0.00017773782019503415
step: 350, loss: 0.00010516717884456739
step: 360, loss: 0.0008023025002330542
step: 370, loss: 7.66499069868587e-05
step: 380, loss: 0.0001124608243117109
step: 390, loss: 5.900010000914335e-05
step: 400, loss: 4.513257590588182e-05
step: 410, loss: 5.1390230510151014e-05
step: 420, loss: 0.0013187191216275096
step: 430, loss: 4.477156471693888e-05
step: 440, loss: 0.0005505033768713474
step: 450, loss: 0.00015697008348070085
step: 460, loss: 8.883116242941469e-05
step: 470, loss: 0.0008919524261727929
step: 480, loss: 5.481762491399422e-05
step: 490, loss: 6.202561780810356e-05
step: 500, loss: 0.0011051902547478676
step: 510, loss: 0.00030143692856654525
step: 520, loss: 0.07254870235919952
step: 530, loss: 0.0001302394812228158
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9197559831065226, f1=0.9108910891089109,
best_f1=0.9202566452795601
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.120293306186795e-05
step: 10, loss: 2.7342024623067118e-05
step: 20, loss: 0.012114305049180984
step: 30, loss: 6.062367174308747e-05
step: 40, loss: 3.0703056836500764e-05

```

step: 50, loss: 6.106304499553517e-05  
step: 60, loss: 9.116841101786122e-05  
step: 70, loss: 3.913948603440076e-05  
step: 80, loss: 3.540377292665653e-05  
step: 90, loss: 0.00015008659102022648  
step: 100, loss: 6.839093839516863e-05  
step: 110, loss: 4.934440949000418e-05  
step: 120, loss: 5.1011022151215e-05  
step: 130, loss: 0.000191405342775397  
step: 140, loss: 5.414237966760993e-05  
step: 150, loss: 3.685274350573309e-05  
step: 160, loss: 0.00020044459961354733  
step: 170, loss: 0.00011058824748033658  
step: 180, loss: 5.380571019486524e-05  
step: 190, loss: 9.123808558797464e-05  
step: 200, loss: 9.801182750379667e-05  
step: 210, loss: 0.0008258800953626633  
step: 220, loss: 3.212171941413544e-05  
step: 230, loss: 0.00023648746719118208  
step: 240, loss: 8.570565114496276e-05  
step: 250, loss: 0.048791710287332535  
step: 260, loss: 4.649687252822332e-05  
step: 270, loss: 0.00015278247883543372  
step: 280, loss: 2.560620487201959e-05  
step: 290, loss: 3.972359263570979e-05  
step: 300, loss: 5.5287022405536845e-05  
step: 310, loss: 4.0372295188717544e-05  
step: 320, loss: 0.000500575581099838  
step: 330, loss: 7.937884947750717e-05  
step: 340, loss: 4.586985596688464e-05  
step: 350, loss: 0.057149194180965424  
step: 360, loss: 4.744059333461337e-05  
step: 370, loss: 6.042971654096618e-05  
step: 380, loss: 0.00045491266064345837  
step: 390, loss: 2.705240694922395e-05  
step: 400, loss: 3.6058900150237605e-05  
step: 410, loss: 3.817655306193046e-05  
step: 420, loss: 0.00025551352882757783  
step: 430, loss: 7.193874625954777e-05  
step: 440, loss: 0.00020757205493282527  
step: 450, loss: 6.532552652060986e-05  
step: 460, loss: 2.3371952920570038e-05  
step: 470, loss: 0.014348138123750687  
step: 480, loss: 4.6699715312570333e-05  
step: 490, loss: 0.0003600795171223581  
step: 500, loss: 6.181978824315593e-05  
step: 510, loss: 0.0003730329917743802  
step: 520, loss: 2.6288545996067114e-05

```

step: 530, loss: 2.9659839128726162e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9248716752216519, f1=0.9142053445850915,
best_f1=0.9202566452795601
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.108011125936173e-05
step: 10, loss: 3.577621464501135e-05
step: 20, loss: 2.1259831555653363e-05
step: 30, loss: 4.7083416575333104e-05
step: 40, loss: 0.0002755089371930808
step: 50, loss: 0.0001397719606757164
step: 60, loss: 2.735789166763425e-05
step: 70, loss: 3.5127592127537355e-05
step: 80, loss: 3.3757369237719104e-05
step: 90, loss: 4.1129387682303786e-05
step: 100, loss: 4.909970084554516e-05
step: 110, loss: 3.1127223337534815e-05
step: 120, loss: 0.00027499886346049607
step: 130, loss: 3.0512570447172038e-05
step: 140, loss: 5.886959843337536e-05
step: 150, loss: 0.014089208096265793
step: 160, loss: 3.55441115971189e-05
step: 170, loss: 0.0002846585994120687
step: 180, loss: 6.349890463752672e-05
step: 190, loss: 7.967792043928057e-05
step: 200, loss: 3.6677531170425937e-05
step: 210, loss: 7.965540862642229e-05
step: 220, loss: 0.060039374977350235
step: 230, loss: 0.00013460867921821773
step: 240, loss: 6.928336370037869e-05
step: 250, loss: 3.7102625356055796e-05
step: 260, loss: 2.068990761472378e-05
step: 270, loss: 0.00015551302931271493
step: 280, loss: 9.900752047542483e-05
step: 290, loss: 3.5662866139318794e-05
step: 300, loss: 4.586302748066373e-05
step: 310, loss: 3.0299928766908124e-05
step: 320, loss: 0.0002565890608821064
step: 330, loss: 2.5308854674221948e-05
step: 340, loss: 8.958832768257707e-05
step: 350, loss: 0.0008517267415300012
step: 360, loss: 0.0003558750031515956
step: 370, loss: 4.3083797208964825e-05
step: 380, loss: 6.080484308768064e-05

```

```

step: 390, loss: 5.122297443449497e-05
step: 400, loss: 3.8431971915997565e-05
step: 410, loss: 4.243121657054871e-05
step: 420, loss: 3.09150927932933e-05
step: 430, loss: 7.826666114851832e-05
step: 440, loss: 2.6347473976784386e-05
step: 450, loss: 2.380775913479738e-05
step: 460, loss: 0.00014211623056326061
step: 470, loss: 3.173311415594071e-05
step: 480, loss: 2.8069207473890856e-05
step: 490, loss: 3.9389324228977785e-05
step: 500, loss: 0.00035532042966224253
step: 510, loss: 3.075851054745726e-05
step: 520, loss: 4.8156824050238356e-05
step: 530, loss: 2.6981622795574367e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9202797202797203, f1=0.9096018735362997,
best_f1=0.9202566452795601
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.8675638532149605e-05
step: 10, loss: 0.00013000921171624213
step: 20, loss: 3.281070166849531e-05
step: 30, loss: 4.175838694209233e-05
step: 40, loss: 3.236505654058419e-05
step: 50, loss: 2.3081243853084743e-05
step: 60, loss: 3.359363108756952e-05
step: 70, loss: 0.00014380391803570092
step: 80, loss: 3.3041051210602745e-05
step: 90, loss: 3.313072011223994e-05
step: 100, loss: 0.00013648971798829734
step: 110, loss: 0.0025153853930532932
step: 120, loss: 0.00015427637845277786
step: 130, loss: 0.0001874113513622433
step: 140, loss: 2.1512962121050805e-05
step: 150, loss: 0.00010355965787312016
step: 160, loss: 5.520858394447714e-05
step: 170, loss: 3.6148063372820616e-05
step: 180, loss: 2.7234729714109562e-05
step: 190, loss: 0.00019370792142581195
step: 200, loss: 4.006264134659432e-05
step: 210, loss: 4.1990861063823104e-05
step: 220, loss: 1.8283310055267066e-05
step: 230, loss: 2.5804278266150504e-05
step: 240, loss: 4.201940100756474e-05

```



```

step: 250, loss: 0.0008141763391904533
step: 260, loss: 2.9607313990709372e-05
step: 270, loss: 6.972379924263805e-05
step: 280, loss: 3.225987893529236e-05
step: 290, loss: 0.00011448707664385438
step: 300, loss: 7.829741662135348e-05
step: 310, loss: 5.2939896704629064e-05
step: 320, loss: 3.9143596950452775e-05
step: 330, loss: 3.504202322801575e-05
step: 340, loss: 0.00016027269884943962
step: 350, loss: 3.7671226891689e-05
step: 360, loss: 4.1757153667276725e-05
step: 370, loss: 3.321926487842575e-05
step: 380, loss: 3.918852235074155e-05
step: 390, loss: 2.6054043701151386e-05
step: 400, loss: 2.2258220269577578e-05
step: 410, loss: 0.00012978297309018672
step: 420, loss: 4.276762410881929e-05
step: 430, loss: 2.344648055441212e-05
step: 440, loss: 0.0001681557041592896
step: 450, loss: 0.0004326155176386237
step: 460, loss: 1.9758632333832793e-05
step: 470, loss: 0.0009476160630583763
step: 480, loss: 5.130860154167749e-05
step: 490, loss: 2.7227370082982816e-05
step: 500, loss: 3.4133438020944595e-05
step: 510, loss: 3.1376603146782145e-05
step: 520, loss: 9.131151455221698e-05
step: 530, loss: 0.005940494127571583
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9220657276995305, f1=0.9107227208313651,
best_f1=0.9202566452795601
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:

```

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:15, 378.54it/s]
load_f1 = 0.9250457038391224
real_f1 = 0.9240622140896615
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:10, 413.15it/s]

```

### 3 DK

```

[ ]: # Para ejecutar dk necesita
      #!pip install -r requirements.txt
      !python -m spacy download en_core_web_lg

```

```

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting en_core_web_lg==2.2.5
  Downloading https://github.com/explosion/spacy-
models/releases/download/en_core_web_lg-2.2.5/en_core_web_lg-2.2.5.tar.gz (827.9
MB)
    |                                     | 827.9 MB 2.3 MB/s
Requirement already satisfied: spacy>=2.2.2 in
/usr/local/lib/python3.7/dist-packages (from en_core_web_lg==2.2.5) (2.2.3)
Requirement already satisfied: wasabi<1.1.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (0.10.1)
Requirement already satisfied: plac<1.2.0,>=0.9.6 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.1.3)
Requirement already satisfied: srsly<1.1.0,>=0.1.0 in

```

```

/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.5)
Requirement already satisfied: setuptools in /usr/local/lib/python3.7/dist-
packages (from spacy>=2.2.2->en_core_web_lg==2.2.5) (57.4.0)
Requirement already satisfied: catalogue<1.1.0,>=0.0.7 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.0)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.7)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (2.23.0)
Requirement already satisfied: blis<0.5.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (0.4.1)
Requirement already satisfied: numpy>=1.15.0 in /usr/local/lib/python3.7/dist-
packages (from spacy>=2.2.2->en_core_web_lg==2.2.5) (1.19.2)
Requirement already satisfied: thinc<7.4.0,>=7.3.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (7.3.1)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (2.0.6)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (3.0.6)
Requirement already satisfied: importlib-metadata>=0.20 in
/usr/local/lib/python3.7/dist-packages (from
catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5) (4.12.0)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-
packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5)
(3.8.1)
Requirement already satisfied: typing-extensions>=3.6.4 in
/usr/local/lib/python3.7/dist-packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5)
(4.1.1)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-
packages (from requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5)
(2.10)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (2022.6.15)
Requirement already satisfied: chardet<4,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (3.0.4)
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in

```

```

/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (1.25.4)
Requirement already satisfied: tqdm<5.0.0,>=4.10.0 in
/usr/local/lib/python3.7/dist-packages (from
thinc<7.4.0,>=7.3.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (4.41.0)
Building wheels for collected packages: en-core-web-lg
  Building wheel for en-core-web-lg (setup.py) ... done
  Created wheel for en-core-web-lg: filename=en_core_web_lg-2.2.5-py3-none-
any.whl size=829180942
sha256=628edac8273df4a2ef032a62aef81ed39eeacd8ad622affd45bb9ed99fce77a1
  Stored in directory: /tmp/pip-ephem-wheel-cache-
uwv4n_mo/wheels/11/95/ba/2c36cc368c0bd339b44a791c2c1881a1fb714b78c29a4cb8f5
Successfully built en-core-web-lg
Installing collected packages: en-core-web-lg
Successfully installed en-core-web-lg-2.2.5
  Download and installation successful
You can now load the model via spacy.load('en_core_web_lg')

```

## 3.1 DK STRUCTURED

### 3.1.1 Beer - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/Beer \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm distilbert \
  --fp16 \
  --save_model \
  --dk general \
  --dk product \
  --summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Beer.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/Beer \

```

```

--input_path input/input_small.jsonl \
--output_path output/output_small.jsonl \
--lm distilbert \
--use_gpu \
--fp16 \
--checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

warnings.warn("An input tensor was not cuda.")

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0 /usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`. In PyTorch 1.1.0 and later, you should call them in the opposite order:

`optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

"https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate", UserWarning)

step: 0, loss: 0.864508867263794

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0 /usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.4117647058823529, f1=0.17391304347826086,
best_f1=0.17391304347826086
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.36903831362724304
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 2: dev_f1=0.34615384615384615, f1=0.31250000000000006,
best_f1=0.17391304347826086
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3215128779411316
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.3939393939393939, f1=0.2631578947368421,
best_f1=0.17391304347826086
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31313538551330566
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.39215686274509803, f1=0.2580645161290323,
best_f1=0.17391304347826086
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20460925996303558
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.5263157894736842, f1=0.0, best_f1=0.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20087531208992004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.42857142857142855, f1=0.3529411764705882, best_f1=0.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2941611409187317
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.46428571428571436, f1=0.2222222222222222, best_f1=0.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3018586039543152
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 8: dev_f1=0.5384615384615384, f1=0.2, best_f1=0.2
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15567243099212646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.6086956521739131, f1=0.0, best_f1=0.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2323312759399414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.5384615384615384, f1=0.1, best_f1=0.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21368472278118134
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.5217391304347826, f1=0.0, best_f1=0.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2504350244998932
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.5185185185185186, f1=0.0, best_f1=0.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16113150119781494
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.5, f1=0.30303030303030304, best_f1=0.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13542413711547852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.52, f1=0.2777777777777778, best_f1=0.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2456870973110199
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.52, f1=0.2777777777777778, best_f1=0.0
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True

```



```

master_weights      : True
loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled             : True
opt_level           : 02
cast_model_type     : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
91it [00:00, 50533.78it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
load_f1 = 0.4878048780487805
real_f1 = 0.48
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:09, 450.35it/s]

```

### 3.1.2 DBLP-ACM - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/DBLP-ACM \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm distilbert \
  --fp16 \
  --save_model \
  --dk general \
  --dk product \

```

```

--summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-ACM.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/DBLP-ACM \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm distilbert \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback.

Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
 /usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning:  
 An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0  
 /usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`.

In PyTorch 1.1.0 and later, you should call them in the opposite order:

`optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result

in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

```
"https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
```

```
step: 0, loss: 0.8027087450027466
step: 10, loss: 0.47775232791900635
step: 20, loss: 0.5805861949920654
step: 30, loss: 0.42377567291259766
step: 40, loss: 0.20134428143501282
step: 50, loss: 0.14864949882030487
step: 60, loss: 0.15565593540668488
step: 70, loss: 0.17295363545417786
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

```
step: 80, loss: 0.2418058067560196
step: 90, loss: 0.014180518686771393
step: 100, loss: 0.17055100202560425
step: 110, loss: 0.14376237988471985
step: 120, loss: 0.09340623021125793
step: 130, loss: 0.025947580114006996
step: 140, loss: 0.10355082899332047
step: 150, loss: 0.02839547209441662
step: 160, loss: 0.3827442228794098
step: 170, loss: 0.0029926092829555273
step: 180, loss: 0.024964747950434685
step: 190, loss: 0.0915209949016571
step: 200, loss: 0.011100699193775654
step: 210, loss: 0.0056426040828228
step: 220, loss: 0.00690191937610507
step: 230, loss: 0.0641120970249176
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
```

```
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

```
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
```

```
warnings.warn(SAVE_STATE_WARNING, UserWarning)
```

```
epoch 1: dev_f1=0.970917225950783, f1=0.9682539682539683,
best_f1=0.9682539682539683
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.21203851699829102
step: 10, loss: 0.009237649850547314
step: 20, loss: 0.1111338809132576
step: 30, loss: 0.009858406148850918
step: 40, loss: 0.0030735130421817303
step: 50, loss: 0.008728033863008022
```

```

step: 60, loss: 0.002585824579000473
step: 70, loss: 0.004109535366296768
step: 80, loss: 0.00861904863268137
step: 90, loss: 0.005640764720737934
step: 100, loss: 0.06441716849803925
step: 110, loss: 0.02928880788385868
step: 120, loss: 0.001866044127382338
step: 130, loss: 0.005497826728969812
step: 140, loss: 0.1370697170495987
step: 150, loss: 0.025486260652542114
step: 160, loss: 0.038691017776727676
step: 170, loss: 0.12384070456027985
step: 180, loss: 0.005502015817910433
step: 190, loss: 0.32645875215530396
step: 200, loss: 0.009199102409183979
step: 210, loss: 0.032789405435323715
step: 220, loss: 0.002539274748414755
step: 230, loss: 0.03322397172451019
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9797752808988766, f1=0.9807474518686297,
best_f1=0.9807474518686297
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.040442902594804764
step: 10, loss: 0.004907360766083002
step: 20, loss: 0.007619012147188187
step: 30, loss: 0.034141913056373596
step: 40, loss: 0.03377247229218483
step: 50, loss: 0.0029853126034140587
step: 60, loss: 0.07757868617773056
step: 70, loss: 0.0034029935486614704
step: 80, loss: 0.11043988168239594
step: 90, loss: 0.05044529587030411
step: 100, loss: 0.001110987039282918
step: 110, loss: 0.019576558843255043
step: 120, loss: 0.0015532757388427854
step: 130, loss: 0.0006663501262664795
step: 140, loss: 0.0005686682416126132
step: 150, loss: 0.0007687287870794535
step: 160, loss: 0.0009514492121525109
step: 170, loss: 0.003288449253886938

```

```

step: 180, loss: 0.004156509879976511
step: 190, loss: 0.09873098880052567
step: 200, loss: 0.0020743045024573803
step: 210, loss: 0.004817384760826826
step: 220, loss: 0.0029790415428578854
step: 230, loss: 0.08156289905309677
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.9821428571428571, f1=0.9765363128491621,
best_f1=0.9765363128491621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001515667769126594
step: 10, loss: 0.0013073603622615337
step: 20, loss: 0.005374635569751263
step: 30, loss: 0.0020003931131213903
step: 40, loss: 0.0087656918913126
step: 50, loss: 0.0018703783862292767
step: 60, loss: 0.024671241641044617
step: 70, loss: 0.033530913293361664
step: 80, loss: 0.2195766270160675
step: 90, loss: 0.005450016353279352
step: 100, loss: 0.0013896467862650752
step: 110, loss: 0.004550996236503124
step: 120, loss: 0.03546508774161339
step: 130, loss: 0.0008696969016455114
step: 140, loss: 0.0008047726005315781
step: 150, loss: 0.003097590059041977
step: 160, loss: 0.0004655025841202587
step: 170, loss: 0.015968382358551025
step: 180, loss: 0.10201362520456314
step: 190, loss: 0.007605517748743296
step: 200, loss: 0.017177583649754524
step: 210, loss: 0.00018700867076404393
step: 220, loss: 0.0006176611059345305
step: 230, loss: 0.0005300341290421784
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.

```

```

warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.9830890642615557, f1=0.9784824462061155,
best_f1=0.9784824462061155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00031358766136690974
step: 10, loss: 0.00026461330708116293
step: 20, loss: 0.00027447633328847587
step: 30, loss: 0.00012697683996520936
step: 40, loss: 0.0006416859687305987
step: 50, loss: 0.000437905837316066
step: 60, loss: 0.0007492079166695476
step: 70, loss: 0.0007523017702624202
step: 80, loss: 0.0017130413325503469
step: 90, loss: 0.010175748728215694
step: 100, loss: 0.003027245867997408
step: 110, loss: 0.0005226333742029965
step: 120, loss: 0.05581989511847496
step: 130, loss: 0.031403280794620514
step: 140, loss: 0.0003306713479105383
step: 150, loss: 0.000204217474674806
step: 160, loss: 0.0005302031640894711
step: 170, loss: 0.005009467247873545
step: 180, loss: 0.0005606600316241384
step: 190, loss: 0.0010953888995572925
step: 200, loss: 0.011555579490959644
step: 210, loss: 0.007885627448558807
step: 220, loss: 0.002751477062702179
step: 230, loss: 0.0008495281217619777
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.9809203142536477, f1=0.9797752808988766,
best_f1=0.9784824462061155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015574534190818667
step: 10, loss: 0.003449647221714258
step: 20, loss: 0.0005603000172413886
step: 30, loss: 0.0019535571336746216
step: 40, loss: 0.00029839432681910694
step: 50, loss: 0.0008614910184405744
step: 60, loss: 0.00844099372625351
step: 70, loss: 0.0005334105808287859
step: 80, loss: 0.008626900613307953
step: 90, loss: 0.00018116930732503533

```

```

step: 100, loss: 0.0001989522424992174
step: 110, loss: 0.05807038024067879
step: 120, loss: 0.0006981102051213384
step: 130, loss: 0.015975847840309143
step: 140, loss: 0.007359550334513187
step: 150, loss: 0.024268189445137978
step: 160, loss: 0.016154730692505836
step: 170, loss: 0.0211474746465683
step: 180, loss: 0.002752518281340599
step: 190, loss: 0.002343717496842146
step: 200, loss: 0.060931138694286346
step: 210, loss: 0.007926738820970058
step: 220, loss: 0.0015980566386133432
step: 230, loss: 0.0003508683294057846
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9761634506242906, f1=0.9702517162471395,
best_f1=0.9784824462061155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01734122261404991
step: 10, loss: 0.00024087315250653774
step: 20, loss: 0.00020679907174780965
step: 30, loss: 0.00027100194711238146
step: 40, loss: 0.0008636368438601494
step: 50, loss: 9.4907074526418e-05
step: 60, loss: 0.00018154457211494446
step: 70, loss: 0.0001596343208802864
step: 80, loss: 0.00015208475815597922
step: 90, loss: 0.00025072216521948576
step: 100, loss: 0.16135936975479126
step: 110, loss: 0.006282764486968517
step: 120, loss: 0.0008023616974242032
step: 130, loss: 0.0009772370103746653
step: 140, loss: 0.007397851906716824
step: 150, loss: 0.0023769133258610964
step: 160, loss: 0.0003888520004693419
step: 170, loss: 0.0002354845346417278
step: 180, loss: 0.00027734984178096056
step: 190, loss: 0.0007778870058245957
step: 200, loss: 0.00334483222104609
step: 210, loss: 0.00015119319141376764
step: 220, loss: 0.00025656764046289027
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 230, loss: 0.02345607802271843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9819413092550789, f1=0.9809203142536477,
best_f1=0.9784824462061155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024409092962741852
step: 10, loss: 0.09518780559301376
step: 20, loss: 0.00029612737125717103
step: 30, loss: 0.001387650496326387
step: 40, loss: 0.0004529580764938146
step: 50, loss: 0.00033333757892251015
step: 60, loss: 0.00047497599734924734
step: 70, loss: 0.00029758145683445036
step: 80, loss: 0.004337096121162176
step: 90, loss: 0.00021133912377990782
step: 100, loss: 0.00021167795057408512
step: 110, loss: 0.0023315593134611845
step: 120, loss: 0.00012766664440277964
step: 130, loss: 0.01977662183344364
step: 140, loss: 0.00012488172797020525
step: 150, loss: 0.09683597087860107
step: 160, loss: 0.00021447008475661278
step: 170, loss: 0.0005058913375250995
step: 180, loss: 0.0001871076092356816
step: 190, loss: 0.00012564161443151534
step: 200, loss: 0.0009874518727883697
step: 210, loss: 0.0003754578356165439
step: 220, loss: 0.0877133160829544
step: 230, loss: 0.014668971300125122
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9807909604519773, f1=0.9808342728297633,
best_f1=0.9784824462061155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004870390985161066
step: 10, loss: 0.00046953113633207977
step: 20, loss: 0.003453536657616496
step: 30, loss: 0.0006147256935946643
step: 40, loss: 9.481444431003183e-05
step: 50, loss: 0.002963929669931531
step: 60, loss: 0.0007738087442703545
step: 70, loss: 0.0006661624065600336
step: 80, loss: 0.024682926014065742

```



```

step: 90, loss: 0.01525044348090887
step: 100, loss: 0.00026679321308620274
step: 110, loss: 0.00021909976203460246
step: 120, loss: 0.024368951097130775
step: 130, loss: 0.00017960509285330772
step: 140, loss: 0.00014109353651292622
step: 150, loss: 0.00011977337999269366
step: 160, loss: 0.00019671788322739303
step: 170, loss: 0.0015109590021893382
step: 180, loss: 0.0013268634211272001
step: 190, loss: 0.0020680923480540514
step: 200, loss: 0.0018193175783380866
step: 210, loss: 0.00038059381768107414
step: 220, loss: 0.026692131534218788
step: 230, loss: 0.018052861094474792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9784335981838819, f1=0.9739524348810873,
best_f1=0.9784824462061155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00042553324601612985
step: 10, loss: 0.00014254506095312536
step: 20, loss: 0.0003030018706340343
step: 30, loss: 0.002637571422383189
step: 40, loss: 0.00014710290997754782
step: 50, loss: 0.0008588589844293892
step: 60, loss: 0.1371108889579773
step: 70, loss: 0.003316557966172695
step: 80, loss: 0.0017374767921864986
step: 90, loss: 0.00031095807207748294
step: 100, loss: 0.00028380355797708035
step: 110, loss: 0.001183554413728416
step: 120, loss: 0.010973915457725525
step: 130, loss: 0.0003308762388769537
step: 140, loss: 0.017744647338986397
step: 150, loss: 0.00010345601913286373
step: 160, loss: 0.0003761674743145704
step: 170, loss: 0.001560849486850202
step: 180, loss: 0.0006988142849877477
step: 190, loss: 0.0004513029707595706
step: 200, loss: 0.00015063864702824503
step: 210, loss: 0.00031824124744161963
step: 220, loss: 0.003399936482310295
step: 230, loss: 0.00013652024790644646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9795454545454545, f1=0.9750566893424036,
best_f1=0.9784824462061155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016118938219733536
step: 10, loss: 0.00015822517161723226
step: 20, loss: 5.518809848581441e-05
step: 30, loss: 0.0005433707847259939
step: 40, loss: 6.981810292927548e-05
step: 50, loss: 0.0007976680644787848
step: 60, loss: 0.000590968003962189
step: 70, loss: 0.000593721168115735
step: 80, loss: 0.0003408241318538785
step: 90, loss: 0.00037506973603740335
step: 100, loss: 0.0002853417245205492
step: 110, loss: 0.00023245587362907827
step: 120, loss: 0.00013573824253398925
step: 130, loss: 0.00011651120439637452
step: 140, loss: 8.989433263195679e-05
step: 150, loss: 0.0002734610461629927
step: 160, loss: 0.0002591582015156746
step: 170, loss: 0.02136934921145439
step: 180, loss: 0.00030185788637027144
step: 190, loss: 0.00017984806618187577
step: 200, loss: 9.604805381968617e-05
step: 210, loss: 9.803931607166305e-05
step: 220, loss: 0.00012896591215394437
step: 230, loss: 0.015655241906642914
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9806598407281, f1=0.975, best_f1=0.9784824462061155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016443764616269618
step: 10, loss: 0.0005950410850346088
step: 20, loss: 6.49088469799608e-05
step: 30, loss: 0.00013310030044522136
step: 40, loss: 5.9876052546314895e-05
step: 50, loss: 9.564690117258579e-05
step: 60, loss: 0.00015681386867072433
step: 70, loss: 0.000203799907467328
step: 80, loss: 9.822912397794425e-05
step: 90, loss: 0.0001073779130820185

```

```

step: 100, loss: 7.598272350151092e-05
step: 110, loss: 0.007497565820813179
step: 120, loss: 7.125062256818637e-05
step: 130, loss: 0.007535878103226423
step: 140, loss: 7.079406350385398e-05
step: 150, loss: 2.7987312932964414e-05
step: 160, loss: 0.027703246101737022
step: 170, loss: 0.00031284301076084375
step: 180, loss: 0.00013835125719197094
step: 190, loss: 0.00030192555277608335
step: 200, loss: 0.0022489477414637804
step: 210, loss: 0.00010795599519042298
step: 220, loss: 0.0032318916637450457
step: 230, loss: 0.00013271161878947169
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9785794813979707, f1=0.9739524348810873,
best_f1=0.9784824462061155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.3594292694469914e-05
step: 10, loss: 5.95227065787185e-05
step: 20, loss: 0.00015147196245379746
step: 30, loss: 0.10625459998846054
step: 40, loss: 0.0002560028515290469
step: 50, loss: 0.00010472645226400346
step: 60, loss: 0.0002714447909966111
step: 70, loss: 0.00010493807349121198
step: 80, loss: 5.563700688071549e-05
step: 90, loss: 5.095886444905773e-05
step: 100, loss: 0.0001079530265997164
step: 110, loss: 0.00024100873270072043
step: 120, loss: 8.86727575561963e-05
step: 130, loss: 0.00028464250499382615
step: 140, loss: 0.00010698704863898456
step: 150, loss: 0.0549149252474308
step: 160, loss: 6.939806189620867e-05
step: 170, loss: 7.219245890155435e-05
step: 180, loss: 0.00014795103925280273
step: 190, loss: 0.0022549156565219164
step: 200, loss: 7.143518450902775e-05
step: 210, loss: 0.0001134178601205349
step: 220, loss: 9.088828664971516e-05
step: 230, loss: 9.109938400797546e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9783845278725825, f1=0.9761634506242906,
best_f1=0.9784824462061155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.5449963838327676e-05
step: 10, loss: 4.6303452108986676e-05
step: 20, loss: 0.02157408930361271
step: 30, loss: 0.000169862832990475
step: 40, loss: 6.260363443288952e-05
step: 50, loss: 0.00010730076610343531
step: 60, loss: 6.624820525757968e-05
step: 70, loss: 0.0002280693151988089
step: 80, loss: 0.0002513215004000813
step: 90, loss: 0.0004356818099040538
step: 100, loss: 0.01542113721370697
step: 110, loss: 0.00030338228680193424
step: 120, loss: 0.005897327326238155
step: 130, loss: 0.000274613150395453
step: 140, loss: 7.6532996899914e-05
step: 150, loss: 0.0001826713705668226
step: 160, loss: 4.332956814323552e-05
step: 170, loss: 5.287992098601535e-05
step: 180, loss: 0.000533730664756149
step: 190, loss: 0.012027909979224205
step: 200, loss: 5.158816929906607e-05
step: 210, loss: 0.000223400944378227
step: 220, loss: 0.00010181299148825929
step: 230, loss: 4.209335020277649e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9784335981838819, f1=0.9785310734463276,
best_f1=0.9784824462061155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.010982906445861e-05
step: 10, loss: 0.0004398087621666491
step: 20, loss: 8.375351171707734e-05
step: 30, loss: 8.439866360276937e-05
step: 40, loss: 9.074493573280051e-05
step: 50, loss: 9.591955313226208e-05
step: 60, loss: 0.0002927514724433422
step: 70, loss: 0.00010340510925743729
step: 80, loss: 0.00017610947543289512
step: 90, loss: 3.669594298116863e-05

```

```

step: 100, loss: 0.00012341792171355337
step: 110, loss: 7.925696991151199e-05
step: 120, loss: 6.931678944965824e-05
step: 130, loss: 7.258672849275172e-05
step: 140, loss: 0.0005722349742427468
step: 150, loss: 9.409854101249948e-05
step: 160, loss: 0.0006529152160510421
step: 170, loss: 4.250648635206744e-05
step: 180, loss: 6.799463153583929e-05
step: 190, loss: 0.01762833073735237
step: 200, loss: 7.38371309125796e-05
step: 210, loss: 0.024906925857067108
step: 220, loss: 0.00017279433086514473
step: 230, loss: 0.000142288175993599
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9794988610478361, f1=0.9761092150170648,
best_f1=0.9784824462061155
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
2473it [00:07, 342.50it/s]
load_f1 = 0.9808342728297633
real_f1 = 0.9796839729119639
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:10, 404.41it/s]

```

### 3.1.3 DBLP-GoogleScholar - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
    --task Structured/DBLP-GoogleScholar \
    --batch_size 32 \
    --max_len 256 \
    --lr 3e-5 \
    --n_epochs 15 \
    --finetuning \
    --lm distilbert \
    --fp16 \
    --save_model \
    --dk general \
    --dk product \
    --summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-GoogleScholar.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
    --task Structured/DBLP-GoogleScholar \
    --input_path input/input_small.jsonl \
    --output_path output/output_small.jsonl \
    --lm distilbert \
    --use_gpu \
    --fp16 \
    --checkpoint_path checkpoints/

```

Selected optimization level O2: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0  
/usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`. In PyTorch 1.1.0 and later, you should call them in the opposite order: `optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

"<https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>", UserWarning)

step: 0, loss: 0.7921016216278076

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

step: 10, loss: 0.46184447407722473

step: 20, loss: 0.5019192695617676

step: 30, loss: 0.43488964438438416

step: 40, loss: 0.41425621509552

step: 50, loss: 0.27349036931991577

step: 60, loss: 0.16618990898132324

step: 70, loss: 0.14050573110580444

step: 80, loss: 0.40009626746177673

step: 90, loss: 0.17500647902488708

step: 100, loss: 0.40700581669807434

step: 110, loss: 0.09561439603567123

step: 120, loss: 0.06038340553641319

step: 130, loss: 0.028967970982193947

```

step: 140, loss: 0.2698914408683777
step: 150, loss: 0.02988274022936821
step: 160, loss: 0.12800170481204987
step: 170, loss: 0.3367164433002472
step: 180, loss: 0.09594625234603882
step: 190, loss: 0.08380863815546036
step: 200, loss: 0.1320362538099289
step: 210, loss: 0.15461398661136627
step: 220, loss: 0.10790625214576721
step: 230, loss: 0.21254827082157135
step: 240, loss: 0.13938605785369873
step: 250, loss: 0.08706869184970856
step: 260, loss: 0.07878605276346207
step: 270, loss: 0.07069223374128342
step: 280, loss: 0.17868541181087494
step: 290, loss: 0.11176461726427078
step: 300, loss: 0.2715456783771515
step: 310, loss: 0.122930146753788
step: 320, loss: 0.09393538534641266
step: 330, loss: 0.14324606955051422
step: 340, loss: 0.3293571174144745
step: 350, loss: 0.06740254908800125
step: 360, loss: 0.10032669454813004
step: 370, loss: 0.14933542907238007
step: 380, loss: 0.17316897213459015
step: 390, loss: 0.05529358237981796
step: 400, loss: 0.05733678489923477
step: 410, loss: 0.07304227352142334
step: 420, loss: 0.11619564145803452
step: 430, loss: 0.13987299799919128
step: 440, loss: 0.10048066824674606
step: 450, loss: 0.042650677263736725
step: 460, loss: 0.04191875830292702
step: 470, loss: 0.21982280910015106
step: 480, loss: 0.21434828639030457
step: 490, loss: 0.06291724741458893
step: 500, loss: 0.043723054230213165
step: 510, loss: 0.09988827258348465
step: 520, loss: 0.22775176167488098
step: 530, loss: 0.13795994222164154
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9179882193022202, f1=0.91, best_f1=0.91

```



/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning:  
An input tensor was not cuda.

warnings.warn("An input tensor was not cuda.")

step: 0, loss: 0.17127788066864014  
step: 10, loss: 0.10635996609926224  
step: 20, loss: 0.13518227636814117  
step: 30, loss: 0.054772283881902695  
step: 40, loss: 0.013320268131792545  
step: 50, loss: 0.15508820116519928  
step: 60, loss: 0.13983213901519775  
step: 70, loss: 0.12379474937915802  
step: 80, loss: 0.04440716281533241  
step: 90, loss: 0.01043107733130455  
step: 100, loss: 0.3180800974369049  
step: 110, loss: 0.033428288996219635  
step: 120, loss: 0.11492059379816055  
step: 130, loss: 0.05099617689847946  
step: 140, loss: 0.021737800911068916  
step: 150, loss: 0.019725695252418518  
step: 160, loss: 0.10975378751754761  
step: 170, loss: 0.11268968135118484  
step: 180, loss: 0.028756126761436462  
step: 190, loss: 0.09860002249479294  
step: 200, loss: 0.039367321878671646  
step: 210, loss: 0.06133859232068062  
step: 220, loss: 0.15273630619049072  
step: 230, loss: 0.028457339853048325  
step: 240, loss: 0.10356401652097702  
step: 250, loss: 0.1225760206580162  
step: 260, loss: 0.010101377964019775  
step: 270, loss: 0.057919103652238846  
step: 280, loss: 0.17931467294692993  
step: 290, loss: 0.04136335849761963  
step: 300, loss: 0.01655610278248787  
step: 310, loss: 0.03944021090865135  
step: 320, loss: 0.06590545177459717  
step: 330, loss: 0.01770888827741146  
step: 340, loss: 0.009565232321619987  
step: 350, loss: 0.06033366546034813  
step: 360, loss: 0.07275828719139099  
step: 370, loss: 0.013109073974192142  
step: 380, loss: 0.05938290059566498  
step: 390, loss: 0.09180500358343124  
step: 400, loss: 0.07348459213972092  
step: 410, loss: 0.0016145650297403336  
step: 420, loss: 0.10245644301176071  
step: 430, loss: 0.013277724385261536  
step: 440, loss: 0.026992209255695343

```

step: 450, loss: 0.006668176036328077
step: 460, loss: 0.1700887680053711
step: 470, loss: 0.023673532530665398
step: 480, loss: 0.15093834698200226
step: 490, loss: 0.03607534244656563
step: 500, loss: 0.10582329332828522
step: 510, loss: 0.020668532699346542
step: 520, loss: 0.05193024128675461
step: 530, loss: 0.16182847321033478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9216674301420064, f1=0.9158186864014802,
best_f1=0.9158186864014802
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.058160919696092606
step: 10, loss: 0.09664043039083481
step: 20, loss: 0.15532638132572174
step: 30, loss: 0.2326727658510208
step: 40, loss: 0.01913376897573471
step: 50, loss: 0.058162156492471695
step: 60, loss: 0.01825306937098503
step: 70, loss: 0.05781982094049454
step: 80, loss: 0.00290886964648962
step: 90, loss: 0.04920629411935806
step: 100, loss: 0.03685023635625839
step: 110, loss: 0.03623858466744423
step: 120, loss: 0.00427889684215188
step: 130, loss: 0.017206652089953423
step: 140, loss: 0.06609083712100983
step: 150, loss: 0.06688589602708817
step: 160, loss: 0.04452668875455856
step: 170, loss: 0.004775198642164469
step: 180, loss: 0.017438694834709167
step: 190, loss: 0.0037737684324383736
step: 200, loss: 0.05472279712557793
step: 210, loss: 0.1099051833152771
step: 220, loss: 0.08814762532711029
step: 230, loss: 0.08475053310394287
step: 240, loss: 0.008322147652506828
step: 250, loss: 0.04718540608882904
step: 260, loss: 0.012123813852667809

```

```

step: 270, loss: 0.003513199742883444
step: 280, loss: 0.0175639558583498
step: 290, loss: 0.03736562281847
step: 300, loss: 0.0609741136431694
step: 310, loss: 0.08244992047548294
step: 320, loss: 0.06666862219572067
step: 330, loss: 0.0020278829615563154
step: 340, loss: 0.006671383511275053
step: 350, loss: 0.03237699717283249
step: 360, loss: 0.008255413733422756
step: 370, loss: 0.008316132239997387
step: 380, loss: 0.010314357466995716
step: 390, loss: 0.0027358538936823606
step: 400, loss: 0.021916069090366364
step: 410, loss: 0.021689578890800476
step: 420, loss: 0.06952062249183655
step: 430, loss: 0.037948012351989746
step: 440, loss: 0.08111492544412613
step: 450, loss: 0.023120615631341934
step: 460, loss: 0.11839897930622101
step: 470, loss: 0.03230699151754379
step: 480, loss: 0.0074845473282039165
step: 490, loss: 0.010144667699933052
step: 500, loss: 0.053853150457143784
step: 510, loss: 0.016302328556776047
step: 520, loss: 0.014450793154537678
step: 530, loss: 0.01311841793358326
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.9226441631504922, f1=0.9170182841068917,
best_f1=0.9170182841068917
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04156414791941643
step: 10, loss: 0.013844417408108711
step: 20, loss: 0.024770446121692657
step: 30, loss: 0.004569909069687128
step: 40, loss: 0.0035619745030999184
step: 50, loss: 0.019583994522690773
step: 60, loss: 0.0005384100950323045
step: 70, loss: 0.0009907449129968882
step: 80, loss: 0.06894800066947937

```

```
step: 90, loss: 0.007802681997418404
step: 100, loss: 0.17983824014663696
step: 110, loss: 0.010279507376253605
step: 120, loss: 0.0020459063816815615
step: 130, loss: 0.01779174618422985
step: 140, loss: 0.008790913969278336
step: 150, loss: 0.0006865797913633287
step: 160, loss: 0.013292254880070686
step: 170, loss: 0.0013269480550661683
step: 180, loss: 0.009350502863526344
step: 190, loss: 0.0008545786840841174
step: 200, loss: 0.001845218357630074
step: 210, loss: 0.0031513068825006485
step: 220, loss: 0.005637628957629204
step: 230, loss: 0.31933337450027466
step: 240, loss: 0.018262693658471107
step: 250, loss: 0.00601212540641427
step: 260, loss: 0.11486473679542542
step: 270, loss: 0.02957773022353649
step: 280, loss: 0.030179694294929504
step: 290, loss: 0.009465510956943035
step: 300, loss: 0.04373861104249954
step: 310, loss: 0.004673502873629332
step: 320, loss: 0.010809024795889854
step: 330, loss: 0.0662277489900589
step: 340, loss: 0.007289989851415157
step: 350, loss: 0.012024459429085255
step: 360, loss: 0.012567815370857716
step: 370, loss: 0.01754557341337204
step: 380, loss: 0.002160562900826335
step: 390, loss: 0.04032214730978012
step: 400, loss: 0.040573131293058395
step: 410, loss: 0.019119592383503914
step: 420, loss: 0.01596732623875141
step: 430, loss: 0.012444213964045048
step: 440, loss: 0.12356394529342651
step: 450, loss: 0.018578480929136276
step: 460, loss: 0.0037140718195587397
step: 470, loss: 0.024084167554974556
step: 480, loss: 0.0022252460476011038
step: 490, loss: 0.11133690178394318
step: 500, loss: 0.0040721348486840725
step: 510, loss: 0.09831880778074265
step: 520, loss: 0.009395795874297619
step: 530, loss: 0.0016249024774879217
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
epoch 4: dev_f1=0.9215419501133787, f1=0.9242009132420091,  
best_f1=0.9170182841068917  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.010936276987195015  
step: 10, loss: 0.005906932987272739  
step: 20, loss: 0.0012221173383295536  
step: 30, loss: 0.0031166623812168837  
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0  
step: 40, loss: 0.013505475595593452  
step: 50, loss: 0.0022555377800017595  
step: 60, loss: 0.001890268875285983  
step: 70, loss: 0.0022846765350550413  
step: 80, loss: 0.002718997886404395  
step: 90, loss: 0.02076699584722519  
step: 100, loss: 0.0006934602279216051  
step: 110, loss: 0.007242480758577585  
step: 120, loss: 0.02046279050409794  
step: 130, loss: 0.012895924039185047  
step: 140, loss: 0.0015954236732795835  
step: 150, loss: 0.0011331770801916718  
step: 160, loss: 0.18369460105895996  
step: 170, loss: 0.03584326058626175  
step: 180, loss: 0.005105284042656422  
step: 190, loss: 0.001948544173501432  
step: 200, loss: 0.001184436958283186  
step: 210, loss: 0.0009287067223340273  
step: 220, loss: 0.0004033211444038898  
step: 230, loss: 0.0015082743484526873  
step: 240, loss: 0.0014375445898622274  
step: 250, loss: 0.00034224646515212953  
step: 260, loss: 0.02083892561495304  
step: 270, loss: 0.07532145828008652  
step: 280, loss: 0.0262575875967741  
step: 290, loss: 0.0013624762650579214  
step: 300, loss: 0.013475733809173107  
step: 310, loss: 0.0018571477849036455  
step: 320, loss: 0.014505390077829361  
step: 330, loss: 0.002438289811834693  
step: 340, loss: 0.0035651009529829025  
step: 350, loss: 0.005015362519770861  
step: 360, loss: 0.005543679930269718  
step: 370, loss: 0.002028243849053979  
step: 380, loss: 0.0291043221950531  
step: 390, loss: 0.017796354368329048  
step: 400, loss: 0.03946788236498833  
step: 410, loss: 0.001996629638597369
```

```
step: 420, loss: 0.010360944084823132
step: 430, loss: 0.013371940702199936
step: 440, loss: 0.0022308698389679193
step: 450, loss: 0.011143922805786133
step: 460, loss: 0.0012021393049508333
step: 470, loss: 0.016841795295476913
step: 480, loss: 0.0023283534683287144
step: 490, loss: 0.009444154798984528
step: 500, loss: 0.0038216400425881147
step: 510, loss: 0.02969970926642418
step: 520, loss: 0.000647136359475553
step: 530, loss: 0.14971333742141724
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.9185667752442996, f1=0.9131652661064426,
best_f1=0.9170182841068917
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03460901975631714
step: 10, loss: 0.0024605439975857735
step: 20, loss: 0.0004642726562451571
step: 30, loss: 0.010058019310235977
step: 40, loss: 0.00042471964843571186
step: 50, loss: 0.0493030920624733
step: 60, loss: 0.0003262750105932355
step: 70, loss: 0.02682407759130001
step: 80, loss: 0.005082120653241873
step: 90, loss: 0.005380577407777309
step: 100, loss: 0.12737149000167847
step: 110, loss: 0.0074609131552278996
step: 120, loss: 0.0016060869675129652
step: 130, loss: 0.0007293631206266582
step: 140, loss: 0.0012115570716559887
step: 150, loss: 0.000774892105255276
step: 160, loss: 0.0006575302686542273
step: 170, loss: 0.0002586951304692775
step: 180, loss: 0.018521372228860855
step: 190, loss: 0.0009023069869726896
step: 200, loss: 0.00019630468159448355
step: 210, loss: 0.014785200357437134
step: 220, loss: 0.0002481045085005462
step: 230, loss: 0.00029194995295256376
step: 240, loss: 0.0010591194732114673
step: 250, loss: 0.003167003160342574
step: 260, loss: 0.0016940083587542176
step: 270, loss: 0.0009476570412516594
```

```

step: 280, loss: 0.06969983875751495
step: 290, loss: 0.009563501924276352
step: 300, loss: 0.09077955037355423
step: 310, loss: 0.011117074638605118
step: 320, loss: 0.004576216451823711
step: 330, loss: 0.031203964725136757
step: 340, loss: 0.00173303228802979
step: 350, loss: 0.045697521418333054
step: 360, loss: 0.0030426234006881714
step: 370, loss: 0.002311257179826498
step: 380, loss: 0.014844609424471855
step: 390, loss: 0.05466717109084129
step: 400, loss: 0.0004971984308212996
step: 410, loss: 0.0006555630825459957
step: 420, loss: 0.03251693397760391
step: 430, loss: 0.0017314390279352665
step: 440, loss: 0.0012036653934046626
step: 450, loss: 0.004843938164412975
step: 460, loss: 0.001227941713295877
step: 470, loss: 0.11409938335418701
step: 480, loss: 0.006445588078349829
step: 490, loss: 0.0006879052380099893
step: 500, loss: 0.004262206144630909
step: 510, loss: 0.0007646107696928084
step: 520, loss: 0.0014277020236477256
step: 530, loss: 0.006183690391480923
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9200367647058824, f1=0.9271461716937355,
best_f1=0.9170182841068917
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0029403751250356436
step: 10, loss: 0.010518483817577362
step: 20, loss: 0.009161471389234066
step: 30, loss: 0.013128281570971012
step: 40, loss: 0.011202890425920486
step: 50, loss: 0.16467976570129395
step: 60, loss: 0.002446119673550129
step: 70, loss: 0.02900642529129982
step: 80, loss: 0.04135047271847725
step: 90, loss: 0.00016717131074983627
step: 100, loss: 0.002305482281371951
step: 110, loss: 0.00030166073702275753
step: 120, loss: 0.0006422910955734551
step: 130, loss: 0.00028616617782972753

```

```

step: 140, loss: 0.0014633891405537724
step: 150, loss: 0.0009605824016034603
step: 160, loss: 0.0004677215765696019
step: 170, loss: 0.00015230414282996207
step: 180, loss: 0.0022642682306468487
step: 190, loss: 0.0010006084339693189
step: 200, loss: 0.0017499157693237066
step: 210, loss: 0.002180263167247176
step: 220, loss: 0.00029079525847919285
step: 230, loss: 0.0005695324507541955
step: 240, loss: 0.001981376903131604
step: 250, loss: 0.000821467547211796
step: 260, loss: 0.009648056700825691
step: 270, loss: 0.007993191480636597
step: 280, loss: 0.06770522147417068
step: 290, loss: 0.004225908312946558
step: 300, loss: 0.00100092520006001
step: 310, loss: 0.0010384940542280674
step: 320, loss: 0.04492638632655144
step: 330, loss: 0.0004614343633875251
step: 340, loss: 0.012164467014372349
step: 350, loss: 0.00021293910685926676
step: 360, loss: 0.0039846086874604225
step: 370, loss: 0.001886844402179122
step: 380, loss: 0.0014553203945979476
step: 390, loss: 0.00023665899061597884
step: 400, loss: 0.00013503099035006016
step: 410, loss: 0.001044721924699843
step: 420, loss: 0.0004117018252145499
step: 430, loss: 8.43318339320831e-05
step: 440, loss: 0.004007863346487284
step: 450, loss: 0.029459884390234947
step: 460, loss: 0.007848529145121574
step: 470, loss: 0.013635914772748947
step: 480, loss: 0.00039978057611733675
step: 490, loss: 0.00370510620996356
step: 500, loss: 0.0006548418314196169
step: 510, loss: 0.001265977043658495
step: 520, loss: 0.00012014662934234366
step: 530, loss: 0.00017717211449053138
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.9267161410018553, f1=0.9230769230769231,

```



```
best_f1=0.9230769230769231
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021009903866797686
step: 10, loss: 0.011362073943018913
step: 20, loss: 0.01586630381643772
step: 30, loss: 0.006283361930400133
step: 40, loss: 0.005335715599358082
step: 50, loss: 0.0008080959087237716
step: 60, loss: 0.0001488614216214046
step: 70, loss: 0.009106829762458801
step: 80, loss: 0.0007653485517948866
step: 90, loss: 0.00047896758769638836
step: 100, loss: 0.0062920390628278255
step: 110, loss: 0.0019927078392356634
step: 120, loss: 0.00802635308355093
step: 130, loss: 0.0006326173315756023
step: 140, loss: 5.905274520046078e-05
step: 150, loss: 0.0017285416834056377
step: 160, loss: 0.00021100671438034624
step: 170, loss: 0.00415092334151268
step: 180, loss: 0.011963564902544022
step: 190, loss: 0.0026148941833525896
step: 200, loss: 0.001406994997523725
step: 210, loss: 0.0005317020113579929
step: 220, loss: 0.0019562451634556055
step: 230, loss: 0.00030543087632395327
step: 240, loss: 7.168593583628535e-05
step: 250, loss: 0.0004258175613358617
step: 260, loss: 0.028329122811555862
step: 270, loss: 0.0002051872288575396
step: 280, loss: 0.00012978643644601107
step: 290, loss: 0.002580855041742325
step: 300, loss: 6.819625559728593e-05
step: 310, loss: 0.0494125634431839
step: 320, loss: 0.00033584985067136586
step: 330, loss: 0.004589825868606567
step: 340, loss: 0.0005698919994756579
step: 350, loss: 0.001651303842663765
step: 360, loss: 0.007064089644700289
step: 370, loss: 0.0007855719304643571
step: 380, loss: 0.0017524637514725327
step: 390, loss: 0.0005731424316763878
step: 400, loss: 0.1910981833934784
step: 410, loss: 0.00193442078307271
step: 420, loss: 0.009059304371476173
step: 430, loss: 0.004132612608373165
```

```

step: 440, loss: 0.0037480525206774473
step: 450, loss: 0.01865566335618496
step: 460, loss: 0.0010602897964417934
step: 470, loss: 0.0015304480912163854
step: 480, loss: 0.0010348842479288578
step: 490, loss: 0.00026769822579808533
step: 500, loss: 0.0024074921384453773
step: 510, loss: 0.0010662104468792677
step: 520, loss: 0.0017502441769465804
step: 530, loss: 0.0005246854270808399
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9257884972170687, f1=0.9207232267037554,
best_f1=0.9230769230769231
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005406438140198588
step: 10, loss: 0.00031986419344320893
step: 20, loss: 0.0006626032409258187
step: 30, loss: 0.0003917560388799757
step: 40, loss: 0.028479639440774918
step: 50, loss: 0.0002400970843154937
step: 60, loss: 0.00012955292186234146
step: 70, loss: 0.09917990118265152
step: 80, loss: 0.0002102823491441086
step: 90, loss: 0.00030314282048493624
step: 100, loss: 0.0004973169998265803
step: 110, loss: 0.00014298150199465454
step: 120, loss: 0.00012486206833273172
step: 130, loss: 0.0004956095945090055
step: 140, loss: 0.0010889563709497452
step: 150, loss: 0.00011717445158865303
step: 160, loss: 0.00010496582399355248
step: 170, loss: 0.0010499836644157767
step: 180, loss: 0.00020207525813020766
step: 190, loss: 0.001662509166635573
step: 200, loss: 0.00014887258294038475
step: 210, loss: 7.00223827152513e-05
step: 220, loss: 0.00029426245600916445
step: 230, loss: 0.0007376831490546465
step: 240, loss: 0.00034768361365422606
step: 250, loss: 0.00014975968224462122
step: 260, loss: 0.00030667943065054715
step: 270, loss: 0.06706909835338593
step: 280, loss: 0.00010302988084731624
step: 290, loss: 0.0015732464380562305

```

```

step: 300, loss: 0.0005208688089624047
step: 310, loss: 0.00011712778359651566
step: 320, loss: 0.0004962066886946559
step: 330, loss: 0.05119708180427551
step: 340, loss: 0.0006414445233531296
step: 350, loss: 0.0003907368809450418
step: 360, loss: 0.027962351217865944
step: 370, loss: 0.00011877884389832616
step: 380, loss: 9.353468340123072e-05
step: 390, loss: 0.00036812140024267137
step: 400, loss: 0.0049141161143779755
step: 410, loss: 0.0007340788724832237
step: 420, loss: 0.0024529858492314816
step: 430, loss: 0.0001049996935762465
step: 440, loss: 0.004208889789879322
step: 450, loss: 0.012682424858212471
step: 460, loss: 0.00011136042303405702
step: 470, loss: 0.0003147793177049607
step: 480, loss: 0.004251122009009123
step: 490, loss: 0.00010531941370572895
step: 500, loss: 0.0009842676809057593
step: 510, loss: 0.03993050381541252
step: 520, loss: 0.012507512234151363
step: 530, loss: 0.00023426278494298458
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9193473193473194, f1=0.9210649229332087,
best_f1=0.9230769230769231
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002935508673544973
step: 10, loss: 3.849037966574542e-05
step: 20, loss: 0.00011824471584986895
step: 30, loss: 6.073158510844223e-05
step: 40, loss: 0.00014875635679345578
step: 50, loss: 0.0006003683665767312
step: 60, loss: 0.00022007788356859237
step: 70, loss: 0.001475551282055676
step: 80, loss: 0.0009070579544641078
step: 90, loss: 0.00022239414101932198
step: 100, loss: 0.00029556258232332766
step: 110, loss: 0.0007027289248071611
step: 120, loss: 0.00015606233500875533
step: 130, loss: 7.242374704219401e-05
step: 140, loss: 6.401755672413856e-05
step: 150, loss: 8.194534893846139e-05

```

```

step: 160, loss: 9.702030365588143e-05
step: 170, loss: 5.3577230573864654e-05
step: 180, loss: 0.0017691809916868806
step: 190, loss: 0.00012818112736567855
step: 200, loss: 0.0001942871604114771
step: 210, loss: 0.0036603587213903666
step: 220, loss: 0.00013640908582601696
step: 230, loss: 0.00014424078108277172
step: 240, loss: 6.154868606245145e-05
step: 250, loss: 0.00017713583656586707
step: 260, loss: 0.007444562390446663
step: 270, loss: 0.01386196631938219
step: 280, loss: 0.0004231715574860573
step: 290, loss: 9.707026765681803e-05
step: 300, loss: 0.0004644592700060457
step: 310, loss: 6.256744381971657e-05
step: 320, loss: 0.00010740004654508084
step: 330, loss: 0.00014373623707797378
step: 340, loss: 0.001372288097627461
step: 350, loss: 4.6008841309230775e-05
step: 360, loss: 5.869123197044246e-05
step: 370, loss: 0.00024702041991986334
step: 380, loss: 0.0001261633588001132
step: 390, loss: 8.925847942009568e-05
step: 400, loss: 6.60597943351604e-05
step: 410, loss: 0.00010641496919561177
step: 420, loss: 0.00031239085365086794
step: 430, loss: 0.0014051491161808372
step: 440, loss: 5.4795418691355735e-05
step: 450, loss: 0.00028817635029554367
step: 460, loss: 0.0005193035467527807
step: 470, loss: 3.475571429589763e-05
step: 480, loss: 0.00010067656694445759
step: 490, loss: 0.05185152590274811
step: 500, loss: 0.0006816617678850889
step: 510, loss: 0.0005635542329400778
step: 520, loss: 0.0001025121528073214
step: 530, loss: 3.98877018596977e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9257037378864791, f1=0.9216589861751152,
best_f1=0.9230769230769231
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016560668882448226
step: 10, loss: 0.0007817143923602998

```

step: 20, loss: 3.582402132451534e-05  
step: 30, loss: 0.00976276583969593  
step: 40, loss: 0.00012993853306397796  
step: 50, loss: 0.0014520407421514392  
step: 60, loss: 0.014636009931564331  
step: 70, loss: 0.00014767034735996276  
step: 80, loss: 0.00012394736404530704  
step: 90, loss: 8.738195174373686e-05  
step: 100, loss: 9.119318565353751e-05  
step: 110, loss: 5.6417626183247194e-05  
step: 120, loss: 6.978416786296293e-05  
step: 130, loss: 8.136230462696403e-05  
step: 140, loss: 0.0011229496449232101  
step: 150, loss: 0.00043396122055128217  
step: 160, loss: 0.00017848974675871432  
step: 170, loss: 0.0002382572420174256  
step: 180, loss: 0.00012298216461203992  
step: 190, loss: 0.005915234796702862  
step: 200, loss: 0.0001942720264196396  
step: 210, loss: 4.292442463338375e-05  
step: 220, loss: 0.002291053533554077  
step: 230, loss: 0.00026130929472856224  
step: 240, loss: 5.119433262734674e-05  
step: 250, loss: 0.0007145598065108061  
step: 260, loss: 3.591408676584251e-05  
step: 270, loss: 0.017989616841077805  
step: 280, loss: 0.005391902290284634  
step: 290, loss: 0.016355426982045174  
step: 300, loss: 0.0005014306516386569  
step: 310, loss: 0.003222063183784485  
step: 320, loss: 0.010809818282723427  
step: 330, loss: 0.0014143233420327306  
step: 340, loss: 5.7944049331126735e-05  
step: 350, loss: 0.0003355058142915368  
step: 360, loss: 0.0005688292440026999  
step: 370, loss: 5.216056160861626e-05  
step: 380, loss: 2.1554140403168276e-05  
step: 390, loss: 0.002439748030155897  
step: 400, loss: 7.939386705402285e-05  
step: 410, loss: 0.0004093546303920448  
step: 420, loss: 0.00031134465825743973  
step: 430, loss: 6.467939965659752e-05  
step: 440, loss: 3.8719026633771136e-05  
step: 450, loss: 3.190222923876718e-05  
step: 460, loss: 0.0023146155290305614  
step: 470, loss: 0.00029785544029437006  
step: 480, loss: 5.619028888759203e-05  
step: 490, loss: 0.0005320192431099713

```
step: 500, loss: 0.00018783489940688014
step: 510, loss: 3.569823093130253e-05
step: 520, loss: 0.00012836787209380418
step: 530, loss: 3.805098822340369e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9244320815948077, f1=0.9215143120960295,
best_f1=0.9230769230769231
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.937612149864435e-05
step: 10, loss: 0.0001437021273886785
step: 20, loss: 5.189640069147572e-05
step: 30, loss: 4.151642497163266e-05
step: 40, loss: 9.67142332228832e-05
step: 50, loss: 0.0037438026629388332
step: 60, loss: 6.17006080574356e-05
step: 70, loss: 4.991427340428345e-05
step: 80, loss: 0.0017498424276709557
step: 90, loss: 4.51586238341406e-05
step: 100, loss: 0.0003890722873620689
step: 110, loss: 2.10660946322605e-05
step: 120, loss: 8.409078873228282e-05
step: 130, loss: 0.00012577806774061173
step: 140, loss: 3.423781890887767e-05
step: 150, loss: 0.0002905841975007206
step: 160, loss: 4.165031714364886e-05
step: 170, loss: 0.0002460586547385901
step: 180, loss: 6.471870437962934e-05
step: 190, loss: 3.82975340471603e-05
step: 200, loss: 5.0557202484924346e-05
step: 210, loss: 0.00029105492285452783
step: 220, loss: 0.0007116731139831245
step: 230, loss: 0.0010187254520133138
step: 240, loss: 2.9733962946920656e-05
step: 250, loss: 9.032808884512633e-05
step: 260, loss: 0.00043627340346574783
step: 270, loss: 4.740770236821845e-05
step: 280, loss: 4.0833499951986596e-05
step: 290, loss: 0.00016051268903538585
step: 300, loss: 0.00024382720584981143
step: 310, loss: 3.546338120941073e-05
step: 320, loss: 7.650833867955953e-05
step: 330, loss: 0.0001301371812587604
step: 340, loss: 8.133728988468647e-05
step: 350, loss: 0.0038665765896439552
```

```

step: 360, loss: 0.0002052744966931641
step: 370, loss: 0.004552791360765696
step: 380, loss: 4.8780686483951285e-05
step: 390, loss: 2.5782132070162334e-05
step: 400, loss: 2.6441437512403354e-05
step: 410, loss: 0.006252744235098362
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 420, loss: 0.00039279102929867804
step: 430, loss: 6.0825725086033344e-05
step: 440, loss: 5.4624269978376105e-05
step: 450, loss: 3.554422073648311e-05
step: 460, loss: 0.00842429231852293
step: 470, loss: 3.364102667546831e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 480, loss: 0.00010507558909012005
step: 490, loss: 0.00037155207246541977
step: 500, loss: 0.006637658458203077
step: 510, loss: 4.414180148160085e-05
step: 520, loss: 0.011877504177391529
step: 530, loss: 0.00012498673459049314
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9215777262180975, f1=0.9258406264394289,
best_f1=0.9230769230769231
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.555104947532527e-05
step: 10, loss: 8.498524402966723e-05
step: 20, loss: 0.031040096655488014
step: 30, loss: 0.0002696998417377472
step: 40, loss: 3.0747676646569744e-05
step: 50, loss: 0.00017715885769575834
step: 60, loss: 3.860233846353367e-05
step: 70, loss: 4.120716766919941e-05
step: 80, loss: 2.380404475843534e-05
step: 90, loss: 0.00019945891108363867
step: 100, loss: 1.9229601093684323e-05
step: 110, loss: 0.00027862045681104064
step: 120, loss: 4.2353971366537735e-05
step: 130, loss: 0.0010239470284432173
step: 140, loss: 4.811080725630745e-05
step: 150, loss: 7.487546099582687e-05
step: 160, loss: 9.695124754216522e-05
step: 170, loss: 0.0008880391833372414
step: 180, loss: 0.0001685553725110367
step: 190, loss: 4.259237175574526e-05

```

```

step: 200, loss: 0.0012004519812762737
step: 210, loss: 0.006376231089234352
step: 220, loss: 4.614957288140431e-05
step: 230, loss: 3.13840682792943e-05
step: 240, loss: 5.9089310525450855e-05
step: 250, loss: 0.06008656322956085
step: 260, loss: 3.314495552331209e-05
step: 270, loss: 3.9134025428211316e-05
step: 280, loss: 1.7426660633645952e-05
step: 290, loss: 4.848061871598475e-05
step: 300, loss: 0.00012729992158710957
step: 310, loss: 2.6605021048453636e-05
step: 320, loss: 6.459317228291184e-05
step: 330, loss: 0.00024084356846287847
step: 340, loss: 7.459452172042802e-05
step: 350, loss: 0.07481695711612701
step: 360, loss: 2.808412136801053e-05
step: 370, loss: 5.031282489653677e-05
step: 380, loss: 0.000786254764534533
step: 390, loss: 0.00020828295964747667
step: 400, loss: 5.13479026267305e-05
step: 410, loss: 4.437530515133403e-05
step: 420, loss: 3.6233839637134224e-05
step: 430, loss: 4.8330373829230666e-05
step: 440, loss: 6.12495859968476e-05
step: 450, loss: 4.374986019684002e-05
step: 460, loss: 2.5278417524532415e-05
step: 470, loss: 0.003765987465158105
step: 480, loss: 0.0013217696687206626
step: 490, loss: 6.521045725094154e-05
step: 500, loss: 0.011790630407631397
step: 510, loss: 3.517133154673502e-05
step: 520, loss: 4.5912784116808325e-05
step: 530, loss: 3.250891313655302e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9235048678720444, f1=0.9208699676075892,
best_f1=0.9230769230769231
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.749600364244543e-05
step: 10, loss: 4.79130721942056e-05
step: 20, loss: 2.4515386030543596e-05
step: 30, loss: 6.478399882325903e-05
step: 40, loss: 2.831832716765348e-05
step: 50, loss: 0.00010766465857159346

```



step: 60, loss: 2.5670271497801878e-05  
step: 70, loss: 0.00014771844143979251  
step: 80, loss: 3.8694161048624665e-05  
step: 90, loss: 1.9557386622182094e-05  
step: 100, loss: 3.933327388949692e-05  
step: 110, loss: 1.7765611119102687e-05  
step: 120, loss: 1.976160638150759e-05  
step: 130, loss: 2.2198250007932074e-05  
step: 140, loss: 5.5221349612111226e-05  
step: 150, loss: 0.00025086975074373186  
step: 160, loss: 0.0012846096651628613  
step: 170, loss: 8.723066275706515e-05  
step: 180, loss: 8.064429130172357e-05  
step: 190, loss: 5.491622869158164e-05  
step: 200, loss: 3.215170363546349e-05  
step: 210, loss: 4.5077184040565044e-05  
step: 220, loss: 2.778973794193007e-05  
step: 230, loss: 0.0004945889231748879  
step: 240, loss: 2.926842716988176e-05  
step: 250, loss: 2.804679206747096e-05  
step: 260, loss: 1.8599861505208537e-05  
step: 270, loss: 0.0036963869351893663  
step: 280, loss: 6.974704592721537e-05  
step: 290, loss: 9.11457245820202e-05  
step: 300, loss: 3.598228795453906e-05  
step: 310, loss: 1.9926141249015927e-05  
step: 320, loss: 2.942124592664186e-05  
step: 330, loss: 0.00025191326858475804  
step: 340, loss: 5.0959992222487926e-05  
step: 350, loss: 0.00016798068827483803  
step: 360, loss: 0.004061372485011816  
step: 370, loss: 2.3271324607776478e-05  
step: 380, loss: 0.0003121166955679655  
step: 390, loss: 3.14697390422225e-05  
step: 400, loss: 3.845665924018249e-05  
step: 410, loss: 2.7771071472670883e-05  
step: 420, loss: 2.0645140466513112e-05  
step: 430, loss: 4.63496326119639e-05  
step: 440, loss: 1.5765210264362395e-05  
step: 450, loss: 2.7100461011286825e-05  
step: 460, loss: 9.153942664852366e-05  
step: 470, loss: 3.1711344490759075e-05  
step: 480, loss: 3.655194814200513e-05  
step: 490, loss: 2.8362508601276204e-05  
step: 500, loss: 5.517569661606103e-05  
step: 510, loss: 0.0001246204337803647  
step: 520, loss: 5.410812082118355e-05  
step: 530, loss: 2.6419051209813915e-05

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9223616922361693, f1=0.9183387774148392,
best_f1=0.9230769230769231
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.8749422199325636e-05
step: 10, loss: 6.11280047451146e-05
step: 20, loss: 4.974638795829378e-05
step: 30, loss: 8.953803626354784e-05
step: 40, loss: 1.885329584183637e-05
step: 50, loss: 0.00013026755186729133
step: 60, loss: 2.546054201957304e-05
step: 70, loss: 0.0003990682598669082
step: 80, loss: 4.674048977904022e-05
step: 90, loss: 1.8313166947336867e-05
step: 100, loss: 3.059305527131073e-05
step: 110, loss: 0.02572869323194027
step: 120, loss: 0.00025791008374653757
step: 130, loss: 1.7478598238085397e-05
step: 140, loss: 1.8219960111309774e-05
step: 150, loss: 7.100404764059931e-05
step: 160, loss: 3.4758126275846735e-05
step: 170, loss: 4.1429142584092915e-05
step: 180, loss: 5.526406312128529e-05
step: 190, loss: 4.862948844674975e-05
step: 200, loss: 6.633792509092018e-05
step: 210, loss: 2.0432697056094185e-05
step: 220, loss: 1.514681844128063e-05
step: 230, loss: 1.9020983017981052e-05
step: 240, loss: 3.225955879315734e-05
step: 250, loss: 4.8016612709034234e-05
step: 260, loss: 0.00013382610632106662
step: 270, loss: 4.0323757275473326e-05
step: 280, loss: 0.00015127113147173077
step: 290, loss: 0.00032042755628935993
step: 300, loss: 5.2467381465248764e-05
step: 310, loss: 4.304758840589784e-05
step: 320, loss: 3.161531640216708e-05
step: 330, loss: 2.860910535673611e-05
step: 340, loss: 4.355008786660619e-05
step: 350, loss: 3.88887056033127e-05
step: 360, loss: 3.700646630022675e-05
step: 370, loss: 1.8797391021507792e-05
step: 380, loss: 3.755322177312337e-05
step: 390, loss: 2.195621345890686e-05
```

```

step: 400, loss: 1.651767161092721e-05
step: 410, loss: 6.326564471237361e-05
step: 420, loss: 4.442303179530427e-05
step: 430, loss: 1.9523931769072078e-05
step: 440, loss: 3.450977601460181e-05
step: 450, loss: 0.00037219058140181005
step: 460, loss: 1.760539089445956e-05
step: 470, loss: 3.266632484155707e-05
step: 480, loss: 5.5414140661014244e-05
step: 490, loss: 5.5867327318992466e-05
step: 500, loss: 8.738099131733179e-05
step: 510, loss: 0.000116054477985017
step: 520, loss: 2.3729475287836976e-05
step: 530, loss: 2.635181590449065e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.924791086350975, f1=0.9212855146716348,
best_f1=0.9230769230769231
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
5742it [00:14, 401.33it/s]
load_f1 = 0.9256198347107438
real_f1 = 0.9250574712643677
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:11, 396.48it/s]

```

### 3.1.4 Amazon-Google - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
    --task Structured/Amazon-Google \
    --batch_size 32 \
    --max_len 256 \
    --lr 3e-5 \
    --n_epochs 15 \
    --finetuning \
    --lm distilbert \
    --fp16 \
    --save_model \
    --dk general \
    --dk product \
    --summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Amazon-Google.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
    --task Structured/Amazon-Google \
    --input_path input/input_small.jsonl \
    --output_path output/output_small.jsonl \
    --lm distilbert \
    --use_gpu \
    --fp16 \
    --checkpoint_path checkpoints/

```

Selected optimization level O2: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback.

Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

warnings.warn("An input tensor was not cuda.")

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0

/usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`.

In PyTorch 1.1.0 and later, you should call them in the opposite order:

`optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

"<https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>", UserWarning)

step: 0, loss: 0.8332000374794006

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

step: 10, loss: 0.05899498984217644

step: 20, loss: 0.37713363766670227

step: 30, loss: 0.3724764585494995

step: 40, loss: 0.5065053105354309

step: 50, loss: 0.3038850724697113

step: 60, loss: 0.36176466941833496

step: 70, loss: 0.25053834915161133

step: 80, loss: 0.27585265040397644

step: 90, loss: 0.3943423926830292

step: 100, loss: 0.16816376149654388

step: 110, loss: 0.2924734354019165

step: 120, loss: 0.2350534200668335

step: 130, loss: 0.2596733272075653

```

step: 140, loss: 0.22821813821792603
step: 150, loss: 0.246950164437294
step: 160, loss: 0.17601479589939117
step: 170, loss: 0.1252184808254242
step: 180, loss: 0.18859846889972687
step: 190, loss: 0.16412979364395142
step: 200, loss: 0.1496993601322174
step: 210, loss: 0.42887184023857117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.5362637362637362, f1=0.5515695067264574,
best_f1=0.5515695067264574
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10252036154270172
step: 10, loss: 0.10689650475978851
step: 20, loss: 0.2865687906742096
step: 30, loss: 0.11907628178596497
step: 40, loss: 0.12014906853437424
step: 50, loss: 0.15847298502922058
step: 60, loss: 0.06004444509744644
step: 70, loss: 0.12382248044013977
step: 80, loss: 0.17131692171096802
step: 90, loss: 0.2772216796875
step: 100, loss: 0.09472307562828064
step: 110, loss: 0.07130321860313416
step: 120, loss: 0.21125774085521698
step: 130, loss: 0.1755211353302002
step: 140, loss: 0.2823278605937958
step: 150, loss: 0.23692181706428528
step: 160, loss: 0.09298555552959442
step: 170, loss: 0.29548215866088867
step: 180, loss: 0.289889931678772
step: 190, loss: 0.17882145941257477
step: 200, loss: 0.19313910603523254
step: 210, loss: 0.16926774382591248
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.

```

```

    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.5633333333333332, f1=0.6115702479338843,
best_f1=0.6115702479338843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18153177201747894
step: 10, loss: 0.2049267739057541
step: 20, loss: 0.2078966349363327
step: 30, loss: 0.055243246257305145
step: 40, loss: 0.1539401412010193
step: 50, loss: 0.09959347546100616
step: 60, loss: 0.16312240064144135
step: 70, loss: 0.08532576262950897
step: 80, loss: 0.050501465797424316
step: 90, loss: 0.1735392063856125
step: 100, loss: 0.02223830856382847
step: 110, loss: 0.09470465779304504
step: 120, loss: 0.24945560097694397
step: 130, loss: 0.08543482422828674
step: 140, loss: 0.23000317811965942
step: 150, loss: 0.20232217013835907
step: 160, loss: 0.0694240853190422
step: 170, loss: 0.24606716632843018
step: 180, loss: 0.0950268879532814
step: 190, loss: 0.30580782890319824
step: 200, loss: 0.1546725630760193
step: 210, loss: 0.23598645627498627
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.5829959514170041, f1=0.610655737704918,
best_f1=0.610655737704918
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06648419052362442
step: 10, loss: 0.053539473563432693
step: 20, loss: 0.09209512174129486
step: 30, loss: 0.03399166837334633
step: 40, loss: 0.04179651290178299
step: 50, loss: 0.15050284564495087
step: 60, loss: 0.07513945549726486
step: 70, loss: 0.0776248350739479

```

```

step: 80, loss: 0.25811782479286194
step: 90, loss: 0.004217071458697319
step: 100, loss: 0.0510534942150116
step: 110, loss: 0.037156201899051666
step: 120, loss: 0.05368805676698685
step: 130, loss: 0.33820784091949463
step: 140, loss: 0.1665174514055252
step: 150, loss: 0.08096343278884888
step: 160, loss: 0.05007496848702431
step: 170, loss: 0.057373158633708954
step: 180, loss: 0.07887563109397888
step: 190, loss: 0.07457747310400009
step: 200, loss: 0.13928665220737457
step: 210, loss: 0.016435865312814713
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.5761316872427984, f1=0.5995893223819302,
best_f1=0.610655737704918
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0321580246090889
step: 10, loss: 0.13385389745235443
step: 20, loss: 0.03560732677578926
step: 30, loss: 0.08327510952949524
step: 40, loss: 0.051120709627866745
step: 50, loss: 0.12763534486293793
step: 60, loss: 0.02691054157912731
step: 70, loss: 0.07410154491662979
step: 80, loss: 0.02733978070318699
step: 90, loss: 0.06171555817127228
step: 100, loss: 0.034581538289785385
step: 110, loss: 0.01562776416540146
step: 120, loss: 0.005667722783982754
step: 130, loss: 0.2494523674249649
step: 140, loss: 0.03774581104516983
step: 150, loss: 0.10296659916639328
step: 160, loss: 0.023445427417755127
step: 170, loss: 0.024733036756515503
step: 180, loss: 0.036347150802612305
step: 190, loss: 0.22166110575199127
step: 200, loss: 0.12654882669448853
step: 210, loss: 0.0191317368298769
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```



UserWarning: Please also save or load the state of the optimizer when saving or loading the scheduler.

```
warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.5846702317290552, f1=0.6150943396226415,
best_f1=0.6150943396226415
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0291585735976696
step: 10, loss: 0.02395271696150303
step: 20, loss: 0.007394063286483288
step: 30, loss: 0.18522433936595917
step: 40, loss: 0.05205398052930832
step: 50, loss: 0.01496437843888998
step: 60, loss: 0.17283624410629272
step: 70, loss: 0.005693039856851101
step: 80, loss: 0.09316253662109375
step: 90, loss: 0.03693666681647301
step: 100, loss: 0.023398132994771004
step: 110, loss: 0.018962662667036057
step: 120, loss: 0.01576702855527401
step: 130, loss: 0.1069507747888565
step: 140, loss: 0.1440514475107193
step: 150, loss: 0.010594670660793781
step: 160, loss: 0.1870998740196228
step: 170, loss: 0.01502098049968481
step: 180, loss: 0.002252978039905429
step: 190, loss: 0.023058103397488594
step: 200, loss: 0.03220093250274658
step: 210, loss: 0.042550355195999146
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.5637583892617449, f1=0.5831435079726652,
best_f1=0.6150943396226415
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.023067273199558258
step: 10, loss: 0.01004704087972641
step: 20, loss: 0.06358347088098526
step: 30, loss: 0.026252616196870804
step: 40, loss: 0.01593274623155594
step: 50, loss: 0.019733548164367676
step: 60, loss: 0.20686446130275726
step: 70, loss: 0.010447736829519272
step: 80, loss: 0.07083357870578766
step: 90, loss: 0.042532987892627716
```

```

step: 100, loss: 0.007073320914059877
step: 110, loss: 0.013796322047710419
step: 120, loss: 0.10007453709840775
step: 130, loss: 0.005665640812367201
step: 140, loss: 0.008332446217536926
step: 150, loss: 0.02587074786424637
step: 160, loss: 0.10208567976951599
step: 170, loss: 0.0028197953943163157
step: 180, loss: 0.0035711778327822685
step: 190, loss: 0.0162588432431221
step: 200, loss: 0.004328278359025717
step: 210, loss: 0.029122455045580864
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.5690021231422505, f1=0.6090712742980561,
best_f1=0.6150943396226415
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01183042861521244
step: 10, loss: 0.001493348041549325
step: 20, loss: 0.010288078337907791
step: 30, loss: 0.008473964408040047
step: 40, loss: 0.052008017897605896
step: 50, loss: 0.013247424736618996
step: 60, loss: 0.08447744697332382
step: 70, loss: 0.003588117193430662
step: 80, loss: 0.009992617182433605
step: 90, loss: 0.014574253931641579
step: 100, loss: 0.0029488771688193083
step: 110, loss: 0.00438427459448576
step: 120, loss: 0.0023622766602784395
step: 130, loss: 0.006507378537207842
step: 140, loss: 0.002870487980544567
step: 150, loss: 0.0034932217095047235
step: 160, loss: 0.021226156502962112
step: 170, loss: 0.0021486449986696243
step: 180, loss: 0.024128269404172897
step: 190, loss: 0.0012549118837341666
step: 200, loss: 0.002381125232204795
step: 210, loss: 0.0621393658220768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.5685071574642127, f1=0.610878661087866,
best_f1=0.6150943396226415
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0075708720833063126
step: 10, loss: 0.10237327963113785
step: 20, loss: 0.018186228349804878
step: 30, loss: 0.015204456634819508
step: 40, loss: 0.06252079457044601
step: 50, loss: 0.0274092685431242
step: 60, loss: 0.007257594261318445
step: 70, loss: 0.0007510184077546
step: 80, loss: 0.003874179208651185
step: 90, loss: 0.04185222089290619
step: 100, loss: 0.01108312327414751
step: 110, loss: 0.00030716569744981825
step: 120, loss: 0.002228714991360903
step: 130, loss: 0.03245381638407707
step: 140, loss: 0.043511297553777695
step: 150, loss: 0.07165762782096863
step: 160, loss: 0.01467171311378479
step: 170, loss: 0.05317365378141403
step: 180, loss: 0.006531591061502695
step: 190, loss: 0.013767192140221596
step: 200, loss: 0.006647057365626097
step: 210, loss: 0.0391959585249424
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.5923076923076923, f1=0.6147859922178989,
best_f1=0.6147859922178989
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004163302946835756
step: 10, loss: 0.001731272554025054
step: 20, loss: 0.01708829402923584
step: 30, loss: 0.0027043211739510298
step: 40, loss: 0.1483653485774994
step: 50, loss: 0.0024617400486022234
step: 60, loss: 0.0008887316798791289
step: 70, loss: 0.058606769889593124
step: 80, loss: 0.00013864837819710374
step: 90, loss: 0.004424539394676685
step: 100, loss: 0.00030157555011101067
step: 110, loss: 0.06175410747528076

```

```

step: 120, loss: 0.0008504677680321038
step: 130, loss: 0.00027246057288721204
step: 140, loss: 0.00022006283688824624
step: 150, loss: 0.009626742452383041
step: 160, loss: 0.002537894994020462
step: 170, loss: 0.016926711425185204
step: 180, loss: 0.0008831611485220492
step: 190, loss: 0.037840116769075394
step: 200, loss: 0.0017312541604042053
step: 210, loss: 0.019282644614577293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.5826446280991735, f1=0.5970772442588728,
best_f1=0.6147859922178989
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010458574630320072
step: 10, loss: 0.009804143570363522
step: 20, loss: 0.009111503139138222
step: 30, loss: 0.0005283310892991722
step: 40, loss: 0.009272260591387749
step: 50, loss: 0.0055655501782894135
step: 60, loss: 0.0012143364874646068
step: 70, loss: 0.005550161469727755
step: 80, loss: 0.0058177681639790535
step: 90, loss: 0.00028837937861680984
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 100, loss: 0.001022211043164134
step: 110, loss: 0.000633574731182307
step: 120, loss: 0.0022774667013436556
step: 130, loss: 0.012907496653497219
step: 140, loss: 0.04640832543373108
step: 150, loss: 0.010352060198783875
step: 160, loss: 0.0003977289889007807
step: 170, loss: 0.04281391203403473
step: 180, loss: 0.004943992476910353
step: 190, loss: 0.016389261931180954
step: 200, loss: 0.0003654791507869959
step: 210, loss: 0.0010761602316051722
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.5434782608695652, f1=0.5884955752212389,
best_f1=0.6147859922178989
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013597391080111265
step: 10, loss: 0.0015415404923260212
step: 20, loss: 0.0012259673094376922
step: 30, loss: 0.0014541485579684377
step: 40, loss: 0.053185105323791504
step: 50, loss: 0.0005343973753042519
step: 60, loss: 0.005308534484356642
step: 70, loss: 0.0001860814227256924
step: 80, loss: 0.0003244078834541142
step: 90, loss: 0.00013752898667007685
step: 100, loss: 0.0006260629743337631
step: 110, loss: 0.0009364110301248729
step: 120, loss: 0.0018435991369187832
step: 130, loss: 0.001633181469514966
step: 140, loss: 0.0005154653335921466
step: 150, loss: 0.0001712999219307676
step: 160, loss: 0.029329344630241394
step: 170, loss: 0.0003193669836036861
step: 180, loss: 0.00040462115430273116
step: 190, loss: 0.0025464019272476435
step: 200, loss: 0.022320285439491272
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 210, loss: 0.0004905729438178241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.5619834710743802, f1=0.6037735849056604,
best_f1=0.6147859922178989
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.486702816095203e-05
step: 10, loss: 0.007848485372960567
step: 20, loss: 0.00011026675929315388
step: 30, loss: 0.002753607230260968
step: 40, loss: 0.00026590420748107135
step: 50, loss: 0.0007042432553134859
step: 60, loss: 0.0002392382302787155
step: 70, loss: 0.014898481778800488
step: 80, loss: 0.01639416441321373
step: 90, loss: 0.02948366105556488
step: 100, loss: 0.0019540961366146803
step: 110, loss: 0.0002017505030380562
step: 120, loss: 0.00022864255879539996
step: 130, loss: 0.00022367388010025024
step: 140, loss: 0.0008639384177513421
step: 150, loss: 0.001049351179972291

```

```

step: 160, loss: 0.00035396398743614554
step: 170, loss: 0.0006423580925911665
step: 180, loss: 0.06539499014616013
step: 190, loss: 0.0002551897778175771
step: 200, loss: 0.0014200808946043253
step: 210, loss: 0.004243583884090185
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.5420560747663551, f1=0.5740740740740742,
best_f1=0.6147859922178989
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012644915841519833
step: 10, loss: 0.0004765461781062186
step: 20, loss: 0.0005716781015507877
step: 30, loss: 0.0004234914085827768
step: 40, loss: 0.0005045298603363335
step: 50, loss: 0.03647211566567421
step: 60, loss: 0.0035101822577416897
step: 70, loss: 0.0015497939893975854
step: 80, loss: 0.015571708790957928
step: 90, loss: 0.003450337564572692
step: 100, loss: 0.00016397798026446253
step: 110, loss: 0.0003140862099826336
step: 120, loss: 0.0024920697323977947
step: 130, loss: 0.00013680300617124885
step: 140, loss: 0.0003180416824761778
step: 150, loss: 0.0012189577100798488
step: 160, loss: 0.0005309020052663982
step: 170, loss: 0.0002461681724525988
step: 180, loss: 0.0005171780940145254
step: 190, loss: 0.00014390924479812384
step: 200, loss: 0.00014153492520563304
step: 210, loss: 0.0005617171991616488
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.5530973451327434, f1=0.59375, best_f1=0.6147859922178989
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0025610020384192467
step: 10, loss: 0.0007727898191660643
step: 20, loss: 0.0045935166999697685
step: 30, loss: 0.00010075730097014457
step: 40, loss: 0.00012867033365182579

```

```

step: 50, loss: 0.00011551350326044485
step: 60, loss: 0.001651204307563603
step: 70, loss: 0.017822381108999252
step: 80, loss: 0.0020322080235928297
step: 90, loss: 0.000732214713934809
step: 100, loss: 0.0003097326843999326
step: 110, loss: 0.00012186863023089245
step: 120, loss: 0.0003331232292111963
step: 130, loss: 0.00036343224928714335
step: 140, loss: 0.0004833223356399685
step: 150, loss: 0.001447716262191534
step: 160, loss: 0.0014703792985528708
step: 170, loss: 0.0010050996206700802
step: 180, loss: 0.0027538856957107782
step: 190, loss: 0.0005324414232745767
step: 200, loss: 0.014062667265534401
step: 210, loss: 0.02295803837478161
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.5511111111111111, f1=0.5919282511210763,
best_f1=0.6147859922178989
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2293it [00:03, 619.33it/s]
load_f1 = 0.5904059040590406
real_f1 = 0.594392523364486
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:10, 405.44it/s]
```

### 3.1.5 Walmart-Amazon - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/Walmart-Amazon \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm distilbert \
  --fp16 \
  --save_model \
  --dk general \
  --dk product \
  --summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Walmart-Amazon.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/Walmart-Amazon \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm distilbert \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/
```



Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0  
/usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`. In PyTorch 1.1.0 and later, you should call them in the opposite order: `optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

```
"https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
```

```
step: 0, loss: 0.854117214679718
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

```
step: 10, loss: 0.15923932194709778
```

```
step: 20, loss: 0.1554960161447525
```

```
step: 30, loss: 0.5079358816146851
```

```
step: 40, loss: 0.2578115463256836
```

```
step: 50, loss: 0.3132960796356201
```

```
step: 60, loss: 0.36682799458503723
```

```
step: 70, loss: 0.18082307279109955
```

```
step: 80, loss: 0.5510019063949585
```

```
step: 90, loss: 0.24685931205749512
```

```
step: 100, loss: 0.22527168691158295
```

```

step: 110, loss: 0.23551130294799805
step: 120, loss: 0.41615909337997437
step: 130, loss: 0.34639331698417664
step: 140, loss: 0.3276326656341553
step: 150, loss: 0.26082852482795715
step: 160, loss: 0.22292587161064148
step: 170, loss: 0.40315309166908264
step: 180, loss: 0.3050349950790405
step: 190, loss: 0.141097754240036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.5409429280397022, f1=0.5635910224438903,
best_f1=0.5635910224438903
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2664334774017334
step: 10, loss: 0.04886301979422569
step: 20, loss: 0.07666199654340744
step: 30, loss: 0.16846869885921478
step: 40, loss: 0.3912588953971863
step: 50, loss: 0.3417065143585205
step: 60, loss: 0.11655840277671814
step: 70, loss: 0.24784541130065918
step: 80, loss: 0.1693841814994812
step: 90, loss: 0.15003027021884918
step: 100, loss: 0.227712020277977
step: 110, loss: 0.21577920019626617
step: 120, loss: 0.2637203335762024
step: 130, loss: 0.15858162939548492
step: 140, loss: 0.22427929937839508
step: 150, loss: 0.12071234732866287
step: 160, loss: 0.08748691529035568
step: 170, loss: 0.3956814706325531
step: 180, loss: 0.1069663017988205
step: 190, loss: 0.14687860012054443
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)

```

```

epoch 2: dev_f1=0.7086614173228346, f1=0.7474226804123711,
best_f1=0.7474226804123711
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18665923178195953
step: 10, loss: 0.18242228031158447
step: 20, loss: 0.12321446090936661
step: 30, loss: 0.03265209496021271
step: 40, loss: 0.13212357461452484
step: 50, loss: 0.22816839814186096
step: 60, loss: 0.04409782961010933
step: 70, loss: 0.10709881782531738
step: 80, loss: 0.3066710829734802
step: 90, loss: 0.05895804613828659
step: 100, loss: 0.10199356079101562
step: 110, loss: 0.2357538640499115
step: 120, loss: 0.02409403957426548
step: 130, loss: 0.030826596543192863
step: 140, loss: 0.12682902812957764
step: 150, loss: 0.15706318616867065
step: 160, loss: 0.09364099055528641
step: 170, loss: 0.04509689658880234
step: 180, loss: 0.03310099616646767
step: 190, loss: 0.14044500887393951
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.6881720430107526, f1=0.7326203208556149,
best_f1=0.7474226804123711
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07783377915620804
step: 10, loss: 0.029988916590809822
step: 20, loss: 0.0777571052312851
step: 30, loss: 0.06776618957519531
step: 40, loss: 0.002240629866719246
step: 50, loss: 0.030601035803556442
step: 60, loss: 0.04924682900309563
step: 70, loss: 0.006184983532875776
step: 80, loss: 0.07730540633201599
step: 90, loss: 0.09876931458711624
step: 100, loss: 0.003139287466183305
step: 110, loss: 0.012310991995036602
step: 120, loss: 0.02319103293120861
step: 130, loss: 0.3077739179134369
step: 140, loss: 0.07089770585298538

```

```

step: 150, loss: 0.009983520023524761
step: 160, loss: 0.02461809292435646
step: 170, loss: 0.015458276495337486
step: 180, loss: 0.14026132225990295
step: 190, loss: 0.04509637504816055
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.7624309392265193, f1=0.7637362637362638,
best_f1=0.7637362637362638
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0663670152425766
step: 10, loss: 0.0021582941990345716
step: 20, loss: 0.022553781047463417
step: 30, loss: 0.00544022535905242
step: 40, loss: 0.05317007377743721
step: 50, loss: 0.09694970399141312
step: 60, loss: 0.007498176768422127
step: 70, loss: 0.002929371315985918
step: 80, loss: 0.00756250973790884
step: 90, loss: 0.014888077974319458
step: 100, loss: 0.19099369645118713
step: 110, loss: 0.05069815367460251
step: 120, loss: 0.0008497989038005471
step: 130, loss: 0.007097130641341209
step: 140, loss: 0.01627410016953945
step: 150, loss: 0.018795939162373543
step: 160, loss: 0.029289109632372856
step: 170, loss: 0.007464525755494833
step: 180, loss: 0.18475137650966644
step: 190, loss: 0.12609748542308807
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.7210526315789473, f1=0.7094972067039106,
best_f1=0.7637362637362638
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005535873118788004
step: 10, loss: 0.002440469339489937
step: 20, loss: 0.008785665966570377

```

```

step: 30, loss: 0.001274281763471663
step: 40, loss: 0.0028618734795600176
step: 50, loss: 0.0042487699538469315
step: 60, loss: 0.00851257611066103
step: 70, loss: 0.06807179003953934
step: 80, loss: 0.12180822342634201
step: 90, loss: 0.012890001758933067
step: 100, loss: 0.0324300080537796
step: 110, loss: 0.0020017167553305626
step: 120, loss: 0.060056693851947784
step: 130, loss: 0.001468735164962709
step: 140, loss: 0.009481638669967651
step: 150, loss: 0.0511198490858078
step: 160, loss: 0.012109251692891121
step: 170, loss: 0.01736764796078205
step: 180, loss: 0.003595164744183421
step: 190, loss: 0.1472776234149933
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.7292225201072386, f1=0.7616438356164383,
best_f1=0.7637362637362638
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014127391623333097
step: 10, loss: 0.010862244293093681
step: 20, loss: 0.0018007357139140368
step: 30, loss: 0.014002089388668537
step: 40, loss: 0.01491131354123354
step: 50, loss: 0.026568932458758354
step: 60, loss: 0.0013009533286094666
step: 70, loss: 0.016338994726538658
step: 80, loss: 0.004313657060265541
step: 90, loss: 0.0008908627787604928
step: 100, loss: 0.001082080532796681
step: 110, loss: 0.003946731798350811
step: 120, loss: 0.020095808431506157
step: 130, loss: 0.0008511688793078065
step: 140, loss: 0.0062874178402125835
step: 150, loss: 0.019932003691792488
step: 160, loss: 0.006267183925956488
step: 170, loss: 0.021985072642564774
step: 180, loss: 0.0020268154330551624
step: 190, loss: 0.0006809891201555729
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

epoch 7: dev_f1=0.7146814404432134, f1=0.7359550561797753,
best_f1=0.7637362637362638
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005782083608210087
step: 10, loss: 0.003445323323830962
step: 20, loss: 0.0025339953135699034
step: 30, loss: 0.0019908673129975796
step: 40, loss: 0.0018738217186182737
step: 50, loss: 0.0009659026982262731
step: 60, loss: 0.0006418109405785799
step: 70, loss: 0.0063029383309185505
step: 80, loss: 0.01370986271649599
step: 90, loss: 0.0021746372804045677
step: 100, loss: 0.040636446326971054
step: 110, loss: 0.05220601707696915
step: 120, loss: 0.0006324058049358428
step: 130, loss: 0.0006633648299612105
step: 140, loss: 0.06333466619253159
step: 150, loss: 0.001014017267152667
step: 160, loss: 0.0015179344918578863
step: 170, loss: 0.000784180301707238
step: 180, loss: 0.0014936479274183512
step: 190, loss: 0.03368143364787102
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.7088607594936709, f1=0.7493403693931397,
best_f1=0.7637362637362638
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013441629707813263
step: 10, loss: 0.00044596882071346045
step: 20, loss: 0.007456391118466854
step: 30, loss: 0.007545981556177139
step: 40, loss: 0.003304888727143407
step: 50, loss: 0.0004569503362290561
step: 60, loss: 0.0018693371675908566
step: 70, loss: 0.010782880708575249
step: 80, loss: 0.008763996884226799
step: 90, loss: 0.020239021629095078
step: 100, loss: 0.09144441783428192
step: 110, loss: 0.004317745566368103
step: 120, loss: 0.024445129558444023
step: 130, loss: 0.006901581771671772
step: 140, loss: 0.000285045854980126

```

```

step: 150, loss: 0.004233512096107006
step: 160, loss: 0.0006385897286236286
step: 170, loss: 0.0008934983052313328
step: 180, loss: 0.016178850084543228
step: 190, loss: 0.010993028059601784
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.7107438016528925, f1=0.7401129943502824,
best_f1=0.7637362637362638
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00032564238063059747
step: 10, loss: 0.0014565030578523874
step: 20, loss: 0.0013612814946100116
step: 30, loss: 0.0007772720418870449
step: 40, loss: 0.0003068945079576224
step: 50, loss: 0.0006163516663946211
step: 60, loss: 0.006230722647160292
step: 70, loss: 0.15378963947296143
step: 80, loss: 0.0004451540589798242
step: 90, loss: 0.000639246020000428
step: 100, loss: 0.0005647215875796974
step: 110, loss: 0.003684774972498417
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 120, loss: 0.011171551421284676
step: 130, loss: 0.0015268032439053059
step: 140, loss: 0.010160006582736969
step: 150, loss: 0.0002691317640710622
step: 160, loss: 0.0004916363977827132
step: 170, loss: 0.0004907610127702355
step: 180, loss: 0.0005178314750082791
step: 190, loss: 0.0008917360100895166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.7257617728531854, f1=0.7541899441340784,
best_f1=0.7637362637362638
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07158631831407547
step: 10, loss: 0.0010496920440346003
step: 20, loss: 0.0019076826283708215
step: 30, loss: 0.0024431655183434486
step: 40, loss: 0.0006737743387930095
step: 50, loss: 0.00045560032594949007

```

```

step: 60, loss: 0.0008951930212788284
step: 70, loss: 0.00021035721874795854
step: 80, loss: 0.0020776866003870964
step: 90, loss: 0.0005898909294046462
step: 100, loss: 0.00025724523584358394
step: 110, loss: 0.00019786991470027715
step: 120, loss: 0.003517532953992486
step: 130, loss: 0.0009305640705861151
step: 140, loss: 0.0004215116787236184
step: 150, loss: 0.0038962378166615963
step: 160, loss: 0.0003369394689798355
step: 170, loss: 0.011934768408536911
step: 180, loss: 0.009795376099646091
step: 190, loss: 0.0875224769115448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.7348066298342542, f1=0.7645429362880887,
best_f1=0.7637362637362638
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021787572768516839
step: 10, loss: 0.0014101167907938361
step: 20, loss: 0.0006614225567318499
step: 30, loss: 0.012472279369831085
step: 40, loss: 0.001560345757752657
step: 50, loss: 0.08108644187450409
step: 60, loss: 0.00023369169502984732
step: 70, loss: 0.031180374324321747
step: 80, loss: 0.0003858616400975734
step: 90, loss: 0.0007299709250219166
step: 100, loss: 0.0007649307954125106
step: 110, loss: 0.00026745814830064774
step: 120, loss: 0.0002270220429636538
step: 130, loss: 0.0002833637408912182
step: 140, loss: 0.0014253281988203526
step: 150, loss: 0.15806329250335693
step: 160, loss: 0.0010200084652751684
step: 170, loss: 0.0006486578495241702
step: 180, loss: 0.0010354521218687296
step: 190, loss: 0.012347063980996609
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.7262872628726287, f1=0.7513812154696132,
best_f1=0.7637362637362638
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```



An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004828210803680122
step: 10, loss: 0.00033887833706103265
step: 20, loss: 0.0005859197117388248
step: 30, loss: 0.001054162043146789
step: 40, loss: 0.00012649426935240626
step: 50, loss: 0.0001888949773274362
step: 60, loss: 0.0006382561405189335
step: 70, loss: 0.0002839845255948603
step: 80, loss: 0.0038462430238723755
step: 90, loss: 0.0007605105638504028
step: 100, loss: 0.00046381857828237116
step: 110, loss: 0.0008806891273707151
step: 120, loss: 0.0006110261310823262
step: 130, loss: 0.0006448653875850141
step: 140, loss: 0.0026535694487392902
step: 150, loss: 0.003949777688831091
step: 160, loss: 0.13331164419651031
step: 170, loss: 0.00019779225112870336
step: 180, loss: 0.004541825037449598
step: 190, loss: 0.0009798926766961813
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.7179487179487181, f1=0.7392550143266476,
best_f1=0.7637362637362638
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0025202736724168062
step: 10, loss: 0.0065927510149776936
step: 20, loss: 0.00023798031907062978
step: 30, loss: 0.0004447057726792991
step: 40, loss: 0.0001611932530067861
step: 50, loss: 0.054348718374967575
step: 60, loss: 0.0018517376156523824
step: 70, loss: 0.00026203409652225673
step: 80, loss: 7.611512410221621e-05
step: 90, loss: 0.0007008436950854957
step: 100, loss: 0.0004825413052458316
step: 110, loss: 0.00037708209129050374
step: 120, loss: 0.000478632835438475
step: 130, loss: 0.0006139928591437638
step: 140, loss: 0.0071039036847651005
step: 150, loss: 0.00034357717959210277
step: 160, loss: 0.00013496512838173658
step: 170, loss: 0.0007761428132653236
```

```

step: 180, loss: 0.00033493811497464776
step: 190, loss: 0.00045902642887085676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.7204610951008645, f1=0.7321428571428571,
best_f1=0.7637362637362638
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021686717809643596
step: 10, loss: 0.0001423386565875262
step: 20, loss: 0.00021377923258114606
step: 30, loss: 0.0021004138980060816
step: 40, loss: 0.0003450769872870296
step: 50, loss: 0.0032866448163986206
step: 60, loss: 0.0006420824793167412
step: 70, loss: 0.0005105925956740975
step: 80, loss: 0.00012954662088304758
step: 90, loss: 0.0006964661297388375
step: 100, loss: 0.001516680233180523
step: 110, loss: 0.00018496332631912082
step: 120, loss: 0.02113933116197586
step: 130, loss: 0.0003735609643626958
step: 140, loss: 0.0012894048122689128
step: 150, loss: 0.0002404217084404081
step: 160, loss: 0.0005506901070475578
step: 170, loss: 0.0007300894940271974
step: 180, loss: 0.000224655304918997
step: 190, loss: 0.01340527180582285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.7219251336898397, f1=0.7576601671309193,
best_f1=0.7637362637362638
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```
enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic
```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

2049it [00:06, 312.26it/s]

load\_f1 = 0.628992628992629

real\_f1 = 0.6213592233009708

0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

4398it [00:10, 404.18it/s]

## 3.2 DK TEXTUAL

### 3.2.1 Abt-Buy - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Textual/Abt-Buy \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm distilbert \
  --fp16 \
  --save_model \
  --dk general \
  --dk product \
  --summarize
```

```

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Abt-Buy.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Textual/Abt-Buy \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm distilbert \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

Selected optimization level O2: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled          : True
opt_level        : O2
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled          : True
opt_level        : O2
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0  
/usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`.

In PyTorch 1.1.0 and later, you should call them in the opposite order:

`optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more

details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

```
"https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
```

```
step: 0, loss: 0.8506894111633301
```

```
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
```

```
step: 10, loss: 0.2135220468044281
```

```
step: 20, loss: 0.15630947053432465
```

```
step: 30, loss: 0.23352302610874176
```

```
step: 40, loss: 0.31723105907440186
```

```
step: 50, loss: 0.3786143958568573
```

```
step: 60, loss: 0.43932342529296875
```

```
step: 70, loss: 0.3105947971343994
```

```
step: 80, loss: 0.2533433437347412
```

```
step: 90, loss: 0.41134291887283325
```

```
step: 100, loss: 0.21838921308517456
```

```
step: 110, loss: 0.1901341825723648
```

```
step: 120, loss: 0.5522593259811401
```

```
step: 130, loss: 0.41744303703308105
```

```
step: 140, loss: 0.4869214594364166
```

```
step: 150, loss: 0.11893206089735031
```

```
step: 160, loss: 0.32969555258750916
```

```
step: 170, loss: 0.22786575555801392
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
```

```
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

```
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
```

```
warnings.warn(SAVE_STATE_WARNING, UserWarning)
```

```
epoch 1: dev_f1=0.5060827250608273, f1=0.46341463414634143,
```

```
best_f1=0.46341463414634143
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.3105718493461609
```

```
step: 10, loss: 0.13564327359199524
```

```
step: 20, loss: 0.34247511625289917
```

```
step: 30, loss: 0.23785558342933655
```

```
step: 40, loss: 0.12280559539794922
```

```
step: 50, loss: 0.20198115706443787
```

```
step: 60, loss: 0.07372632622718811
```

```
step: 70, loss: 0.37684381008148193
```

```
step: 80, loss: 0.15756294131278992
```

```
step: 90, loss: 0.12398412078619003
```

```
step: 100, loss: 0.1355258673429489
```

```
step: 110, loss: 0.2239682376384735
```

```
step: 120, loss: 0.052412182092666626
```

```

step: 130, loss: 0.12549325823783875
step: 140, loss: 0.13983283936977386
step: 150, loss: 0.16520604491233826
step: 160, loss: 0.10426463186740875
step: 170, loss: 0.2570371925830841
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.7218390804597702, f1=0.7651006711409396,
best_f1=0.7651006711409396
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.055817246437072754
step: 10, loss: 0.1056610718369484
step: 20, loss: 0.05794551968574524
step: 30, loss: 0.1813500076532364
step: 40, loss: 0.023736728355288506
step: 50, loss: 0.07560210675001144
step: 60, loss: 0.08588282763957977
step: 70, loss: 0.16277194023132324
step: 80, loss: 0.04444919899106026
step: 90, loss: 0.12542837858200073
step: 100, loss: 0.0630110427737236
step: 110, loss: 0.11270380020141602
step: 120, loss: 0.003336314344778657
step: 130, loss: 0.12135156244039536
step: 140, loss: 0.011418447829782963
step: 150, loss: 0.13441815972328186
step: 160, loss: 0.07764674723148346
step: 170, loss: 0.07223794609308243
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.7373271889400922, f1=0.761904761904762,
best_f1=0.761904761904762
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06337437778711319

```

```

step: 10, loss: 0.085859514772892
step: 20, loss: 0.036675188690423965
step: 30, loss: 0.024731287732720375
step: 40, loss: 0.006724012549966574
step: 50, loss: 0.03441169485449791
step: 60, loss: 0.07603468745946884
step: 70, loss: 0.004698913544416428
step: 80, loss: 0.038709115236997604
step: 90, loss: 0.025987327098846436
step: 100, loss: 0.015192675404250622
step: 110, loss: 0.018931837752461433
step: 120, loss: 0.08713451027870178
step: 130, loss: 0.04512054845690727
step: 140, loss: 0.006106088403612375
step: 150, loss: 0.024047963321208954
step: 160, loss: 0.04197528585791588
step: 170, loss: 0.010843011550605297
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.7570332480818416, f1=0.7892156862745098,
best_f1=0.7892156862745098
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.035949915647506714
step: 10, loss: 0.016252748668193817
step: 20, loss: 0.006910026539117098
step: 30, loss: 0.16719093918800354
step: 40, loss: 0.006092876195907593
step: 50, loss: 0.012223665602505207
step: 60, loss: 0.0009995747823268175
step: 70, loss: 0.013718548230826855
step: 80, loss: 0.011010929010808468
step: 90, loss: 0.014739157631993294
step: 100, loss: 0.009043583646416664
step: 110, loss: 0.014600076712667942
step: 120, loss: 0.015746353194117546
step: 130, loss: 0.03804149851202965
step: 140, loss: 0.11085597425699234
step: 150, loss: 0.010959797538816929
step: 160, loss: 0.010918956249952316
step: 170, loss: 0.08418979495763779
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.7677261613691931, f1=0.8018648018648018,
best_f1=0.8018648018648018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014993281103670597
step: 10, loss: 0.030068127438426018
step: 20, loss: 0.004578997381031513
step: 30, loss: 0.026798823848366737
step: 40, loss: 0.0025199265219271183
step: 50, loss: 0.1278892308473587
step: 60, loss: 0.05629768595099449
step: 70, loss: 0.02479858510196209
step: 80, loss: 0.07613673061132431
step: 90, loss: 0.030319439247250557
step: 100, loss: 0.09110184758901596
step: 110, loss: 0.005494331941008568
step: 120, loss: 0.014760617166757584
step: 130, loss: 0.10017849504947662
step: 140, loss: 0.03770184889435768
step: 150, loss: 0.14702796936035156
step: 160, loss: 0.1229184940457344
step: 170, loss: 0.005439791362732649
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 6: dev_f1=0.7716535433070866, f1=0.8181818181818181,
best_f1=0.8181818181818181
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013364193961024284
step: 10, loss: 0.0010930560529232025
step: 20, loss: 0.0033648863900452852
step: 30, loss: 0.030105778947472572
step: 40, loss: 0.009518676437437534
step: 50, loss: 0.0027784474659711123
step: 60, loss: 0.19306673109531403

```



```

step: 70, loss: 0.007838895544409752
step: 80, loss: 0.00834723375737667
step: 90, loss: 0.004458386451005936
step: 100, loss: 0.006189448293298483
step: 110, loss: 0.0009442321606911719
step: 120, loss: 0.12614865601062775
step: 130, loss: 0.2644879221916199
step: 140, loss: 0.04875869303941727
step: 150, loss: 0.02803792990744114
step: 160, loss: 0.005945282988250256
step: 170, loss: 0.1296393871307373
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.7486910994764399, f1=0.7794871794871795,
best_f1=0.8181818181818181
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05111275985836983
step: 10, loss: 0.009601544588804245
step: 20, loss: 0.0052587538957595825
step: 30, loss: 0.16266514360904694
step: 40, loss: 0.0031065421644598246
step: 50, loss: 0.003450209740549326
step: 60, loss: 0.007699853740632534
step: 70, loss: 0.007284035440534353
step: 80, loss: 0.005335734225809574
step: 90, loss: 0.03593115136027336
step: 100, loss: 0.015710216015577316
step: 110, loss: 0.0050050667487084866
step: 120, loss: 0.0014018026413396
step: 130, loss: 0.009978057816624641
step: 140, loss: 0.0007086314144544303
step: 150, loss: 0.19307070970535278
step: 160, loss: 0.008685468696057796
step: 170, loss: 0.0022071735002100468
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.7539267015706808, f1=0.803970223325062,
best_f1=0.8181818181818181
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007572233211249113
step: 10, loss: 0.01527707651257515
step: 20, loss: 0.008806217461824417

```

```

step: 30, loss: 0.03243261203169823
step: 40, loss: 0.017602061852812767
step: 50, loss: 0.029743226245045662
step: 60, loss: 0.005942041520029306
step: 70, loss: 0.040078289806842804
step: 80, loss: 0.08689075708389282
step: 90, loss: 0.002486914861947298
step: 100, loss: 0.026021761819720268
step: 110, loss: 0.0031446442008018494
step: 120, loss: 0.0026495533529669046
step: 130, loss: 0.001371145248413086
step: 140, loss: 0.0010702324798330665
step: 150, loss: 0.021107086911797523
step: 160, loss: 0.0523393340408802
step: 170, loss: 0.0035413638688623905
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.7728459530026109, f1=0.8030303030303031,
best_f1=0.8030303030303031
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01590081863105297
step: 10, loss: 0.00541788199916482
step: 20, loss: 0.0031539902556687593
step: 30, loss: 0.0007856981828808784
step: 40, loss: 0.002216898137703538
step: 50, loss: 0.004306952469050884
step: 60, loss: 0.0023671737872064114
step: 70, loss: 0.002038220874965191
step: 80, loss: 0.004399446304887533
step: 90, loss: 0.05479157716035843
step: 100, loss: 0.057161927223205566
step: 110, loss: 0.0010916971368715167
step: 120, loss: 0.09896235167980194
step: 130, loss: 0.004555670544505119
step: 140, loss: 0.009451543912291527
step: 150, loss: 0.0013867474626749754
step: 160, loss: 0.003251287154853344
step: 170, loss: 0.0011031070025637746
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

epoch 10: dev_f1=0.7628865979381445, f1=0.8048780487804877,
best_f1=0.8030303030303031
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004984935512766242
step: 10, loss: 0.02719285525381565
step: 20, loss: 0.006472376640886068
step: 30, loss: 0.0006498657166957855
step: 40, loss: 0.0031544885132461786
step: 50, loss: 0.017554929479956627
step: 60, loss: 0.0032775788567960262
step: 70, loss: 0.01956907846033573
step: 80, loss: 0.0006320199463516474
step: 90, loss: 0.010585840791463852
step: 100, loss: 0.0006436984986066818
step: 110, loss: 0.0709424763917923
step: 120, loss: 0.008342712186276913
step: 130, loss: 0.001257249852642417
step: 140, loss: 0.01199992187321186
step: 150, loss: 0.0012410137569531798
step: 160, loss: 0.011371063999831676
step: 170, loss: 0.09885497391223907
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.7661691542288557, f1=0.7865707434052758,
best_f1=0.8030303030303031
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008146951906383038
step: 10, loss: 0.0017436777707189322
step: 20, loss: 0.023275738582015038
step: 30, loss: 0.00045717498869635165
step: 40, loss: 0.007114643231034279
step: 50, loss: 0.00032364611979573965
step: 60, loss: 0.19048787653446198
step: 70, loss: 0.04477931559085846
step: 80, loss: 0.025176484137773514
step: 90, loss: 0.0005215686396695673
step: 100, loss: 0.004517336841672659
step: 110, loss: 0.000840100459754467
step: 120, loss: 0.007331446744501591
step: 130, loss: 0.001280869240872562
step: 140, loss: 0.04296207055449486
step: 150, loss: 0.0015367463929578662
step: 160, loss: 0.0003945153148379177

```

```

step: 170, loss: 0.0014684811467304826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.7680412371134021, f1=0.800982800982801,
best_f1=0.8030303030303031
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003511980175971985
step: 10, loss: 0.02892017364501953
step: 20, loss: 0.010247809812426567
step: 30, loss: 0.00046203332021832466
step: 40, loss: 0.00035093663609586656
step: 50, loss: 0.002536879386752844
step: 60, loss: 0.000560349493753165
step: 70, loss: 0.007730999030172825
step: 80, loss: 0.0002002568362513557
step: 90, loss: 0.00024441335699521005
step: 100, loss: 0.0007361321477219462
step: 110, loss: 0.0010911384597420692
step: 120, loss: 0.009855818934738636
step: 130, loss: 0.0008529723854735494
step: 140, loss: 0.0001659585686866194
step: 150, loss: 0.0030208686366677284
step: 160, loss: 0.00031766542815603316
step: 170, loss: 0.0009377507958561182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.7672634271099745, f1=0.794188861985472,
best_f1=0.8030303030303031
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007246253080666065
step: 10, loss: 0.0004147585423197597
step: 20, loss: 0.018792446702718735
step: 30, loss: 0.0003808830224443227
step: 40, loss: 0.00030682963551953435
step: 50, loss: 0.05201075226068497
step: 60, loss: 0.0027385600842535496
step: 70, loss: 0.0023742893245071173
step: 80, loss: 0.00618096673861146
step: 90, loss: 0.0003428825584705919
step: 100, loss: 0.0005568311898969114
step: 110, loss: 0.0016424572095274925
step: 120, loss: 0.014052823185920715

```

```

step: 130, loss: 0.0010762768797576427
step: 140, loss: 0.12588118016719818
step: 150, loss: 0.00024245009990409017
step: 160, loss: 0.06070968136191368
step: 170, loss: 0.004305923357605934
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.7688442211055277, f1=0.8009478672985781,
best_f1=0.8030303030303031
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000799714180175215
step: 10, loss: 0.00672109192237258
step: 20, loss: 0.003143579699099064
step: 30, loss: 0.0021361953113228083
step: 40, loss: 0.00019681289268191904
step: 50, loss: 0.0024391685146838427
step: 60, loss: 0.00030244624940678477
step: 70, loss: 0.00030979339499026537
step: 80, loss: 0.0005409755394794047
step: 90, loss: 0.0002686214866116643
step: 100, loss: 0.00045173498801887035
step: 110, loss: 0.005763339344412088
step: 120, loss: 0.0010515226749703288
step: 130, loss: 0.0014470595633611083
step: 140, loss: 0.006551641970872879
step: 150, loss: 0.03199392557144165
step: 160, loss: 0.000116430499474518
step: 170, loss: 0.0022385912016034126
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.7696202531645568, f1=0.8009478672985781,
best_f1=0.8030303030303031
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale              : dynamic

```

```

Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
1916it [00:04, 441.62it/s]
load_f1 = 0.718676122931442
real_f1 = 0.7142857142857142
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:10, 423.24it/s]

```

### 3.3 DK DIRTY

#### 3.3.1 DBLP-ACM - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
    --task Dirty/DBLP-ACM \
    --batch_size 32 \
    --max_len 256 \
    --lr 3e-5 \
    --n_epochs 15 \
    --finetuning \
    --lm distilbert \
    --fp16 \
    --save_model \
    --dk general \
    --dk product \

```

```

--summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-ACM.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Dirty/DBLP-ACM \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm distilbert \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0  
/usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`.

In PyTorch 1.1.0 and later, you should call them in the opposite order:

`optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result

in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

```
"https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
```

```
step: 0, loss: 0.7957992553710938
```

```
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
```

```
step: 10, loss: 0.47772330045700073
```

```
step: 20, loss: 0.6143383979797363
```

```
step: 30, loss: 0.4774108827114105
```

```
step: 40, loss: 0.3117230236530304
```

```
step: 50, loss: 0.10543900728225708
```

```
step: 60, loss: 0.12694498896598816
```

```
step: 70, loss: 0.21446040272712708
```

```
step: 80, loss: 0.20464828610420227
```

```
step: 90, loss: 0.04454497620463371
```

```
step: 100, loss: 0.16317196190357208
```

```
step: 110, loss: 0.1399535834789276
```

```
step: 120, loss: 0.03944818302989006
```

```
step: 130, loss: 0.03812428563833237
```

```
step: 140, loss: 0.03193137049674988
```

```
step: 150, loss: 0.17884275317192078
```

```
step: 160, loss: 0.12164454907178879
```

```
step: 170, loss: 0.01217784732580185
```

```
step: 180, loss: 0.020443962886929512
```

```
step: 190, loss: 0.014233501628041267
```

```
step: 200, loss: 0.01986498013138771
```

```
step: 210, loss: 0.089422307908535
```

```
step: 220, loss: 0.0335102453827858
```

```
step: 230, loss: 0.026936113834381104
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
```

```
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

```
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
```

```
warnings.warn(SAVE_STATE_WARNING, UserWarning)
```

```
epoch 1: dev_f1=0.9621380846325166, f1=0.953125, best_f1=0.953125
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.042333245277404785
```

```
step: 10, loss: 0.002697004470974207
```

```
step: 20, loss: 0.0042775413021445274
```

```
step: 30, loss: 0.0038342184852808714
```

```
step: 40, loss: 0.037616848945617676
```

```
step: 50, loss: 0.015857486054301262
```

```
step: 60, loss: 0.06803850829601288
```



```

step: 70, loss: 0.03816203400492668
step: 80, loss: 0.004832550883293152
step: 90, loss: 0.009596399962902069
step: 100, loss: 0.11083617061376572
step: 110, loss: 0.16339993476867676
step: 120, loss: 0.03657948970794678
step: 130, loss: 0.002333698095753789
step: 140, loss: 0.23512211441993713
step: 150, loss: 0.026523500680923462
step: 160, loss: 0.05801302194595337
step: 170, loss: 0.03832922503352165
step: 180, loss: 0.02476239949464798
step: 190, loss: 0.15943437814712524
step: 200, loss: 0.053386472165584564
step: 210, loss: 0.07579068094491959
step: 220, loss: 0.0033729919232428074
step: 230, loss: 0.014256625436246395
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 2: dev_f1=0.9609810479375697, f1=0.958659217877095, best_f1=0.953125
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09650753438472748
step: 10, loss: 0.045140232890844345
step: 20, loss: 0.012958879582583904
step: 30, loss: 0.007367504294961691
step: 40, loss: 0.05620225891470909
step: 50, loss: 0.0060105049051344395
step: 60, loss: 0.09434197098016739
step: 70, loss: 0.024468684569001198
step: 80, loss: 0.0026729870587587357
step: 90, loss: 0.12273160368204117
step: 100, loss: 0.013882509432733059
step: 110, loss: 0.06449474394321442
step: 120, loss: 0.07483965158462524
step: 130, loss: 0.0029826690442860126
step: 140, loss: 0.0036502964794635773
step: 150, loss: 0.009352392517030239
step: 160, loss: 0.012485270388424397
step: 170, loss: 0.002748316153883934
step: 180, loss: 0.002493828535079956
step: 190, loss: 0.007470771204680204
step: 200, loss: 0.009264688938856125
step: 210, loss: 0.02175682783126831
step: 220, loss: 0.08649414032697678
step: 230, loss: 0.14406971633434296

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.9580973952434882, f1=0.9591836734693877, best_f1=0.953125
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02732623554766178
step: 10, loss: 0.019023137167096138
step: 20, loss: 0.0006690881564281881
step: 30, loss: 0.0028479574248194695
step: 40, loss: 0.06441143900156021
step: 50, loss: 0.01412899512797594
step: 60, loss: 0.003918040078133345
step: 70, loss: 0.0038411791902035475
step: 80, loss: 0.15263891220092773
step: 90, loss: 0.0058148582465946674
step: 100, loss: 0.03284689784049988
step: 110, loss: 0.015394356101751328
step: 120, loss: 0.13199327886104584
step: 130, loss: 0.007906600832939148
step: 140, loss: 0.0037081949412822723
step: 150, loss: 0.009229946881532669
step: 160, loss: 0.0008015068015083671
step: 170, loss: 0.0016526564722880721
step: 180, loss: 0.06330586224794388
step: 190, loss: 0.0034153470769524574
step: 200, loss: 0.0050816903822124004
step: 210, loss: 0.007218003273010254
step: 220, loss: 0.0037895822897553444
step: 230, loss: 0.000732439395505935
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.958997722095672, f1=0.958659217877095, best_f1=0.953125
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000938061042688787
step: 10, loss: 0.0016702573047950864
step: 20, loss: 0.0006506367353722453
step: 30, loss: 0.057526011019945145
step: 40, loss: 0.0007002268685027957
step: 50, loss: 0.0010772800305858254
step: 60, loss: 0.0004742850724142045
step: 70, loss: 0.000369666435290128
step: 80, loss: 0.0024140747264027596
step: 90, loss: 0.0016144572291523218

```

```

step: 100, loss: 0.005448036827147007
step: 110, loss: 0.005185987800359726
step: 120, loss: 0.114324189722538
step: 130, loss: 0.05790559947490692
step: 140, loss: 0.0011080756084993482
step: 150, loss: 0.00036076962715014815
step: 160, loss: 0.005335919093340635
step: 170, loss: 0.2270776778459549
step: 180, loss: 0.0011423737742006779
step: 190, loss: 0.002060218248516321
step: 200, loss: 0.0014815643662586808
step: 210, loss: 0.004328752402216196
step: 220, loss: 0.006565635092556477
step: 230, loss: 0.0010849725222215056
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.9645232815964524, f1=0.9552889858233369,
best_f1=0.9552889858233369
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002980594988912344
step: 10, loss: 0.001562353572808206
step: 20, loss: 0.0007959840004332364
step: 30, loss: 0.004264908842742443
step: 40, loss: 0.0007889471598900855
step: 50, loss: 0.0023011204320937395
step: 60, loss: 0.009213738143444061
step: 70, loss: 0.0006395488162524998
step: 80, loss: 0.005099571775645018
step: 90, loss: 0.0005422980757430196
step: 100, loss: 0.004585083574056625
step: 110, loss: 0.05559134483337402
step: 120, loss: 0.0010491639841347933
step: 130, loss: 0.0011931454064324498
step: 140, loss: 0.0011729826219379902
step: 150, loss: 0.0032940686214715242
step: 160, loss: 0.010302621871232986
step: 170, loss: 0.0028553511947393417
step: 180, loss: 0.006051987409591675
step: 190, loss: 0.002167090307921171
step: 200, loss: 0.010494429618120193
step: 210, loss: 0.0005246114451438189

```

```

step: 220, loss: 0.0005148708005435765
step: 230, loss: 0.0029118626844137907
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9599999999999999, f1=0.9572836801752465,
best_f1=0.9552889858233369
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15577492117881775
step: 10, loss: 0.005431598052382469
step: 20, loss: 0.009361917153000832
step: 30, loss: 0.001152852550148964
step: 40, loss: 0.011022534221410751
step: 50, loss: 0.002130882116034627
step: 60, loss: 0.06047307699918747
step: 70, loss: 0.0003907799255102873
step: 80, loss: 0.00013215870421845466
step: 90, loss: 0.0029756471049040556
step: 100, loss: 0.0005297089810483158
step: 110, loss: 0.00037531580892391503
step: 120, loss: 0.0006856085383333266
step: 130, loss: 0.0002552466175984591
step: 140, loss: 0.00014676418504677713
step: 150, loss: 0.006149642169475555
step: 160, loss: 0.039166562259197235
step: 170, loss: 0.001150517025962472
step: 180, loss: 0.003041572170332074
step: 190, loss: 0.004334867931902409
step: 200, loss: 0.024894094094634056
step: 210, loss: 0.003658626927062869
step: 220, loss: 0.0010550342267379165
step: 230, loss: 0.025349359959363937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9569160997732427, f1=0.9605411499436302,
best_f1=0.9552889858233369
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03996271640062332
step: 10, loss: 0.014693900011479855
step: 20, loss: 0.0004778136790264398
step: 30, loss: 0.0001258616684935987
step: 40, loss: 0.00036137300776317716
step: 50, loss: 0.0005751296994276345

```

```

step: 60, loss: 0.0007669867482036352
step: 70, loss: 0.0019258774118497968
step: 80, loss: 0.04611564055085182
step: 90, loss: 0.0011275960132479668
step: 100, loss: 0.0028049801476299763
step: 110, loss: 0.0018750530434772372
step: 120, loss: 0.0012915311381220818
step: 130, loss: 0.004994583781808615
step: 140, loss: 0.0003115964063908905
step: 150, loss: 0.020986229181289673
step: 160, loss: 0.0026630335487425327
step: 170, loss: 0.000621626153588295
step: 180, loss: 0.00336986081674695
step: 190, loss: 0.0008975992095656693
step: 200, loss: 0.013708771206438541
step: 210, loss: 0.006441487930715084
step: 220, loss: 0.0011788565898314118
step: 230, loss: 0.03843802586197853
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9617117117117117, f1=0.9631284916201117,
best_f1=0.9552889858233369
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022862388286739588
step: 10, loss: 0.002053387463092804
step: 20, loss: 0.0006957951700314879
step: 30, loss: 0.00038035912439227104
step: 40, loss: 0.00010909414413617924
step: 50, loss: 0.00015894735406618565
step: 60, loss: 0.000730811501853168
step: 70, loss: 0.00016106908151414245
step: 80, loss: 0.028094196692109108
step: 90, loss: 0.008922893553972244
step: 100, loss: 0.0013951731380075216
step: 110, loss: 0.003005446633324027
step: 120, loss: 0.06435161083936691
step: 130, loss: 9.015923569677398e-05
step: 140, loss: 0.004265074152499437
step: 150, loss: 0.003256687195971608
step: 160, loss: 0.0001306835183640942
step: 170, loss: 6.044240581104532e-05
step: 180, loss: 0.00011873176845256239
step: 190, loss: 7.798149454174563e-05
step: 200, loss: 8.824989345157519e-05
step: 210, loss: 0.0001532736059743911

```

```

step: 220, loss: 0.04875161126255989
step: 230, loss: 0.024921702221035957
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9609810479375697, f1=0.958057395143488,
best_f1=0.9552889858233369
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021734241454396397
step: 10, loss: 0.00011820705549325794
step: 20, loss: 0.00017357786418870091
step: 30, loss: 0.0002485055010765791
step: 40, loss: 7.873190043028444e-05
step: 50, loss: 0.021651655435562134
step: 60, loss: 0.0177475418895483
step: 70, loss: 6.62717066006735e-05
step: 80, loss: 0.00010547285637585446
step: 90, loss: 8.880234236130491e-05
step: 100, loss: 0.0005231058457866311
step: 110, loss: 0.00010001193004427478
step: 120, loss: 0.09374909847974777
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 130, loss: 0.0024161487817764282
step: 140, loss: 0.031608838587999344
step: 150, loss: 0.00014229353109840304
step: 160, loss: 0.0012184480438008904
step: 170, loss: 0.00032162421848624945
step: 180, loss: 0.002230333397164941
step: 190, loss: 0.0012881263392046094
step: 200, loss: 7.994271436473355e-05
step: 210, loss: 5.912629058002494e-05
step: 220, loss: 0.002064715838059783
step: 230, loss: 0.000265736278379336
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 10: dev_f1=0.9659863945578231, f1=0.967305524239008,
best_f1=0.967305524239008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002275791484862566

```

```

step: 10, loss: 0.005102808121591806
step: 20, loss: 4.03170743084047e-05
step: 30, loss: 0.00011095691297668964
step: 40, loss: 0.0036837339866906404
step: 50, loss: 0.0003169045376125723
step: 60, loss: 9.651776781538501e-05
step: 70, loss: 0.005592084024101496
step: 80, loss: 8.706506923772395e-05
step: 90, loss: 6.618037150474265e-05
step: 100, loss: 8.74569668667391e-05
step: 110, loss: 0.0005577576230280101
step: 120, loss: 0.00016160664381459355
step: 130, loss: 4.790056846104562e-05
step: 140, loss: 0.005521059036254883
step: 150, loss: 5.137893094797619e-05
step: 160, loss: 6.587057578144595e-05
step: 170, loss: 0.0033818010706454515
step: 180, loss: 6.923652836121619e-05
step: 190, loss: 7.807630754541606e-05
step: 200, loss: 0.00020148823386989534
step: 210, loss: 5.1961327699245885e-05
step: 220, loss: 7.097821799106896e-05
step: 230, loss: 0.002951416652649641
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9638009049773756, f1=0.967670011148272,
best_f1=0.967305524239008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.243027019081637e-05
step: 10, loss: 0.0072431969456374645
step: 20, loss: 4.946155604557134e-05
step: 30, loss: 0.05057121813297272
step: 40, loss: 5.590009459410794e-05
step: 50, loss: 0.00014184509927872568
step: 60, loss: 0.006478691007941961
step: 70, loss: 5.0325324991717935e-05
step: 80, loss: 0.0002433494955766946
step: 90, loss: 0.0002648528607096523
step: 100, loss: 5.004945705877617e-05
step: 110, loss: 8.484758291160688e-05
step: 120, loss: 0.04857847839593887
step: 130, loss: 0.0003831423819065094
step: 140, loss: 9.101487376028672e-05
step: 150, loss: 3.7515725125558674e-05
step: 160, loss: 0.0030673202127218246

```

```

step: 170, loss: 5.587255145655945e-05
step: 180, loss: 5.8534998970571905e-05
step: 190, loss: 4.933838499709964e-05
step: 200, loss: 0.0062162489630281925
step: 210, loss: 6.642964581260458e-05
step: 220, loss: 0.014526418410241604
step: 230, loss: 6.24424428679049e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9625425652667423, f1=0.9630459126539753,
best_f1=0.967305524239008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.358087557600811e-05
step: 10, loss: 0.00012701874948106706
step: 20, loss: 7.200665277196094e-05
step: 30, loss: 0.005861369427293539
step: 40, loss: 0.0007256789831444621
step: 50, loss: 0.0006415083771571517
step: 60, loss: 0.0001276779657928273
step: 70, loss: 0.00013682892313227057
step: 80, loss: 4.076528057339601e-05
step: 90, loss: 0.0007756931008771062
step: 100, loss: 5.511221388587728e-05
step: 110, loss: 0.00016507507825735956
step: 120, loss: 0.0010165674611926079
step: 130, loss: 0.0002300996711710468
step: 140, loss: 0.0009987051598727703
step: 150, loss: 0.010237851180136204
step: 160, loss: 0.0001401990302838385
step: 170, loss: 0.00011407298006815836
step: 180, loss: 0.00041035006870515645
step: 190, loss: 7.828987145330757e-05
step: 200, loss: 0.00010298829874955118
step: 210, loss: 7.191023178165779e-05
step: 220, loss: 5.836861237185076e-05
step: 230, loss: 5.0615020882105455e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9627118644067796, f1=0.9622222222222223,
best_f1=0.967305524239008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.233792323153466e-05

```



```

step: 10, loss: 2.9943012123112567e-05
step: 20, loss: 0.00013041961938142776
step: 30, loss: 6.573151040356606e-05
step: 40, loss: 3.677410495583899e-05
step: 50, loss: 0.00012105728819733486
step: 60, loss: 3.845433457172476e-05
step: 70, loss: 5.083900032332167e-05
step: 80, loss: 9.728677105158567e-05
step: 90, loss: 0.0005817707278765738
step: 100, loss: 0.002180302981287241
step: 110, loss: 0.00011992120562354103
step: 120, loss: 3.7836165574844927e-05
step: 130, loss: 8.059066021814942e-05
step: 140, loss: 3.843971717287786e-05
step: 150, loss: 0.0005414577317424119
step: 160, loss: 3.0188872187864035e-05
step: 170, loss: 6.171976565383375e-05
step: 180, loss: 0.00012297515058889985
step: 190, loss: 0.0009611977729946375
step: 200, loss: 3.7817517295479774e-05
step: 210, loss: 0.022495398297905922
step: 220, loss: 0.0004020134510938078
step: 230, loss: 2.7912863515666686e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9625425652667423, f1=0.9631284916201117,
best_f1=0.967305524239008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.188965198816732e-05
step: 10, loss: 4.8858604714041576e-05
step: 20, loss: 5.358747876016423e-05
step: 30, loss: 0.0002524868759792298
step: 40, loss: 7.061631913529709e-05
step: 50, loss: 0.00013149582082405686
step: 60, loss: 3.316128277219832e-05
step: 70, loss: 5.809022695757449e-05
step: 80, loss: 0.00411826279014349
step: 90, loss: 4.94135201734025e-05
step: 100, loss: 4.749156505567953e-05
step: 110, loss: 4.754306428367272e-05
step: 120, loss: 5.161065928405151e-05
step: 130, loss: 4.522275412455201e-05
step: 140, loss: 5.038169911131263e-05
step: 150, loss: 7.453245780197904e-05
step: 160, loss: 3.651020597317256e-05

```

```

step: 170, loss: 4.453604560694657e-05
step: 180, loss: 4.621384869096801e-05
step: 190, loss: 0.00037475398858077824
step: 200, loss: 4.642903877538629e-05
step: 210, loss: 0.004270379431545734
step: 220, loss: 0.0057643502950668335
step: 230, loss: 6.781404226785526e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9625425652667423, f1=0.9620535714285715,
best_f1=0.967305524239008
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:06, 354.20it/s]
load_f1 = 0.9624724061810155
real_f1 = 0.96353591160221
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

4398it [00:11, 394.47it/s]

### 3.3.2 DBLP-GoogleScholar - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Dirty/DBLP-GoogleScholar \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm distilbert \
  --fp16 \
  --save_model \
  --dk general \
  --dk product \
  --summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-GoogleScholar/.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Dirty/DBLP-GoogleScholar \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm distilbert \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/
```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

|                       |                 |
|-----------------------|-----------------|
| enabled               | : True          |
| opt_level             | : 02            |
| cast_model_type       | : torch.float16 |
| patch_torch_functions | : False         |
| keep_batchnorm_fp32   | : True          |
| master_weights        | : True          |

```

loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled             : True
opt_level           : 02
cast_model_type     : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.7917529344558716
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4499722421169281
step: 20, loss: 0.5019012689590454
step: 30, loss: 0.45575398206710815
step: 40, loss: 0.44222545623779297
step: 50, loss: 0.34726306796073914
step: 60, loss: 0.2781902253627777
step: 70, loss: 0.24067631363868713
step: 80, loss: 0.08705444633960724
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 90, loss: 0.17545859515666962
step: 100, loss: 0.2990912199020386
step: 110, loss: 0.04907038062810898
step: 120, loss: 0.08135685324668884
step: 130, loss: 0.034470897167921066
step: 140, loss: 0.12919457256793976
step: 150, loss: 0.029204027727246284
step: 160, loss: 0.28859347105026245
step: 170, loss: 0.14345616102218628
step: 180, loss: 0.16292844712734222
step: 190, loss: 0.03868481144309044

```

```

step: 200, loss: 0.08902598172426224
step: 210, loss: 0.07853146642446518
step: 220, loss: 0.13367797434329987
step: 230, loss: 0.15550647675991058
step: 240, loss: 0.12650172412395477
step: 250, loss: 0.08538244664669037
step: 260, loss: 0.03591359779238701
step: 270, loss: 0.023628441616892815
step: 280, loss: 0.16144824028015137
step: 290, loss: 0.15960142016410828
step: 300, loss: 0.14464133977890015
step: 310, loss: 0.060365479439496994
step: 320, loss: 0.08772822469472885
step: 330, loss: 0.12104234099388123
step: 340, loss: 0.3249987065792084
step: 350, loss: 0.04899252951145172
step: 360, loss: 0.08666544407606125
step: 370, loss: 0.16592320799827576
step: 380, loss: 0.2015116810798645
step: 390, loss: 0.04967019334435463
step: 400, loss: 0.025623193010687828
step: 410, loss: 0.04024490341544151
step: 420, loss: 0.03414086624979973
step: 430, loss: 0.03551562875509262
step: 440, loss: 0.25000324845314026
step: 450, loss: 0.014039617031812668
step: 460, loss: 0.03655710816383362
step: 470, loss: 0.1173776313662529
step: 480, loss: 0.2951006293296814
step: 490, loss: 0.07075385004281998
step: 500, loss: 0.011897267773747444
step: 510, loss: 0.13350875675678253
step: 520, loss: 0.07407566905021667
step: 530, loss: 0.054852742701768875
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9182736455463728, f1=0.9064354176175262,
best_f1=0.9064354176175262
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08033130317926407
step: 10, loss: 0.15624871850013733

```

step: 20, loss: 0.13732409477233887  
step: 30, loss: 0.04942582547664642  
step: 40, loss: 0.00994477141648531  
step: 50, loss: 0.05421360954642296  
step: 60, loss: 0.09408455342054367  
step: 70, loss: 0.19034625589847565  
step: 80, loss: 0.0095107052475214  
step: 90, loss: 0.0017152527580037713  
step: 100, loss: 0.28274640440940857  
step: 110, loss: 0.0374947227537632  
step: 120, loss: 0.09393797814846039  
step: 130, loss: 0.014375320635735989  
step: 140, loss: 0.014759835787117481  
step: 150, loss: 0.03469991683959961  
step: 160, loss: 0.030312255024909973  
step: 170, loss: 0.018423397094011307  
step: 180, loss: 0.0035467317793518305  
step: 190, loss: 0.02136509120464325  
step: 200, loss: 0.030378157272934914  
step: 210, loss: 0.0106259286403656  
step: 220, loss: 0.15393908321857452  
step: 230, loss: 0.014874747954308987  
step: 240, loss: 0.15502168238162994  
step: 250, loss: 0.03571043163537979  
step: 260, loss: 0.005211370065808296  
step: 270, loss: 0.12913842499256134  
step: 280, loss: 0.29524165391921997  
step: 290, loss: 0.09261620044708252  
step: 300, loss: 0.045789994299411774  
step: 310, loss: 0.04125944525003433  
step: 320, loss: 0.0652366429567337  
step: 330, loss: 0.0274592787027359  
step: 340, loss: 0.006120861507952213  
step: 350, loss: 0.060070764273405075  
step: 360, loss: 0.03864344581961632  
step: 370, loss: 0.009430796839296818  
step: 380, loss: 0.09812914580106735  
step: 390, loss: 0.018957385793328285  
step: 400, loss: 0.0766274705529213  
step: 410, loss: 0.0007564832922071218  
step: 420, loss: 0.0574035570025444  
step: 430, loss: 0.048873916268348694  
step: 440, loss: 0.006361743435263634  
step: 450, loss: 0.01242192555218935  
step: 460, loss: 0.2363368570804596  
step: 470, loss: 0.024154001846909523  
step: 480, loss: 0.26832520961761475  
step: 490, loss: 0.0527469702064991

```

step: 500, loss: 0.04533298686146736
step: 510, loss: 0.06823064386844635
step: 520, loss: 0.08889459818601608
step: 530, loss: 0.10718919336795807
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9189944134078213, f1=0.9184909175593852,
best_f1=0.9184909175593852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007832692936062813
step: 10, loss: 0.16761791706085205
step: 20, loss: 0.24365679919719696
step: 30, loss: 0.16301023960113525
step: 40, loss: 0.003032359294593334
step: 50, loss: 0.01518929935991764
step: 60, loss: 0.004248172510415316
step: 70, loss: 0.0035499788355082273
step: 80, loss: 0.01012360118329525
step: 90, loss: 0.12520736455917358
step: 100, loss: 0.012372983619570732
step: 110, loss: 0.026564909145236015
step: 120, loss: 0.006981397978961468
step: 130, loss: 0.026556529104709625
step: 140, loss: 0.01306125707924366
step: 150, loss: 0.0021272823214530945
step: 160, loss: 0.012784304097294807
step: 170, loss: 0.012681988999247551
step: 180, loss: 0.07844974845647812
step: 190, loss: 0.029131190851330757
step: 200, loss: 0.00536896288394928
step: 210, loss: 0.04807688295841217
step: 220, loss: 0.0497460775077343
step: 230, loss: 0.045983973890542984
step: 240, loss: 0.012169482186436653
step: 250, loss: 0.02947048656642437
step: 260, loss: 0.0008702423656359315
step: 270, loss: 0.006799647584557533
step: 280, loss: 0.00605900539085269
step: 290, loss: 0.028230052441358566
step: 300, loss: 0.05131062492728233
step: 310, loss: 0.09239981323480606

```

```

step: 320, loss: 0.13254855573177338
step: 330, loss: 0.035859666764736176
step: 340, loss: 0.004744293633848429
step: 350, loss: 0.021043356508016586
step: 360, loss: 0.035903673619031906
step: 370, loss: 0.009340543299913406
step: 380, loss: 0.010402405634522438
step: 390, loss: 0.008607188239693642
step: 400, loss: 0.009059995412826538
step: 410, loss: 0.015701470896601677
step: 420, loss: 0.039330609142780304
step: 430, loss: 0.01002044603228569
step: 440, loss: 0.022206731140613556
step: 450, loss: 0.07180317491292953
step: 460, loss: 0.09032410383224487
step: 470, loss: 0.013088635168969631
step: 480, loss: 0.010878588072955608
step: 490, loss: 0.019043955951929092
step: 500, loss: 0.042227379977703094
step: 510, loss: 0.004806001670658588
step: 520, loss: 0.004070254508405924
step: 530, loss: 0.04437842220067978
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.9269427640763147, f1=0.9221501390176089,
best_f1=0.9221501390176089
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011454515159130096
step: 10, loss: 0.01325647160410881
step: 20, loss: 0.009716342203319073
step: 30, loss: 0.001116773346439004
step: 40, loss: 0.008598573505878448
step: 50, loss: 0.03129206970334053
step: 60, loss: 0.003061844501644373
step: 70, loss: 0.007073762360960245
step: 80, loss: 0.03337430581450462
step: 90, loss: 0.05097058415412903
step: 100, loss: 0.07458028197288513
step: 110, loss: 0.0026171798817813396
step: 120, loss: 0.003095024498179555
step: 130, loss: 0.013629871420562267

```



```

step: 140, loss: 0.0030251313000917435
step: 150, loss: 0.001280423253774643
step: 160, loss: 0.024913107976317406
step: 170, loss: 0.019855953752994537
step: 180, loss: 0.0184530857950449
step: 190, loss: 0.0126066654920578
step: 200, loss: 0.002294112229719758
step: 210, loss: 0.03276445344090462
step: 220, loss: 0.003838414093479514
step: 230, loss: 0.2929476499557495
step: 240, loss: 0.001058831694535911
step: 250, loss: 0.018060456961393356
step: 260, loss: 0.12356267124414444
step: 270, loss: 0.0352565161883831
step: 280, loss: 0.00131131405942142
step: 290, loss: 0.007028382271528244
step: 300, loss: 0.003592869034036994
step: 310, loss: 0.008047333918511868
step: 320, loss: 0.027710992842912674
step: 330, loss: 0.1539042443037033
step: 340, loss: 0.030566494911909103
step: 350, loss: 0.0033140622545033693
step: 360, loss: 0.01332787424325943
step: 370, loss: 0.013303331099450588
step: 380, loss: 0.010631999000906944
step: 390, loss: 0.01651565358042717
step: 400, loss: 0.01115479413419962
step: 410, loss: 0.0045294323936104774
step: 420, loss: 0.0014030407182872295
step: 430, loss: 0.012027019634842873
step: 440, loss: 0.038198087364435196
step: 450, loss: 0.0011164648458361626
step: 460, loss: 0.001422698376700282
step: 470, loss: 0.03602367639541626
step: 480, loss: 0.010127764195203781
step: 490, loss: 0.010132074356079102
step: 500, loss: 0.009888466447591782
step: 510, loss: 0.008213679306209087
step: 520, loss: 0.009913153015077114
step: 530, loss: 0.0006452483939938247
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.9206204379562043, f1=0.9168949771689499,
best_f1=0.9221501390176089
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 0.005238217767328024  
step: 10, loss: 0.0017720552859827876  
step: 20, loss: 0.0014258219162002206  
step: 30, loss: 0.002463794546201825  
step: 40, loss: 0.031669486314058304  
step: 50, loss: 0.005475470330566168  
step: 60, loss: 0.0015059506986290216  
step: 70, loss: 0.016171639785170555  
step: 80, loss: 0.0008112813229672611  
step: 90, loss: 0.024327510967850685  
step: 100, loss: 0.003774257143959403  
step: 110, loss: 0.001390065997838974  
step: 120, loss: 0.002345267217606306  
step: 130, loss: 0.000342809857102111  
step: 140, loss: 0.0007701683789491653  
step: 150, loss: 0.000730461732018739  
step: 160, loss: 0.0018263200763612986  
step: 170, loss: 0.005896510556340218  
step: 180, loss: 0.007383678574115038  
step: 190, loss: 0.0005604136385954916  
step: 200, loss: 0.008251192979514599  
step: 210, loss: 0.002387420739978552  
step: 220, loss: 0.0001540071825729683  
step: 230, loss: 0.0014457606012001634  
step: 240, loss: 0.0007989272708073258  
step: 250, loss: 0.0003312775806989521  
step: 260, loss: 0.07336768507957458  
step: 270, loss: 0.06368277221918106  
step: 280, loss: 0.04545215889811516  
step: 290, loss: 0.012531526386737823  
step: 300, loss: 0.022033318877220154  
step: 310, loss: 0.0008373785531148314  
step: 320, loss: 0.05657932907342911  
step: 330, loss: 0.0008246635552495718  
step: 340, loss: 0.0012280568480491638  
step: 350, loss: 0.004937055055052042  
step: 360, loss: 0.13180841505527496  
step: 370, loss: 0.007786829490214586  
step: 380, loss: 0.038429584354162216  
step: 390, loss: 0.003148091258481145  
step: 400, loss: 0.024468494579195976  
step: 410, loss: 0.00031444805790670216  
step: 420, loss: 0.004963601473718882  
step: 430, loss: 0.0013541732914745808  
step: 440, loss: 0.09822599589824677  
step: 450, loss: 0.10451138764619827  
step: 460, loss: 0.0024539283476769924  
step: 470, loss: 0.002370787551626563

```

step: 480, loss: 0.002180554438382387
step: 490, loss: 0.01810215227305889
step: 500, loss: 0.005616023670881987
step: 510, loss: 0.0012416905956342816
step: 520, loss: 0.0003987833624705672
step: 530, loss: 0.005345695186406374
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.9240037071362373, f1=0.9149232914923291,
best_f1=0.9221501390176089
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014314325526356697
step: 10, loss: 0.0016215663636103272
step: 20, loss: 0.019031718373298645
step: 30, loss: 0.0009221505024470389
step: 40, loss: 0.0008085574372671545
step: 50, loss: 0.0001375928259221837
step: 60, loss: 0.0005662811454385519
step: 70, loss: 0.025014208629727364
step: 80, loss: 0.0006788394530303776
step: 90, loss: 0.0005324773956090212
step: 100, loss: 0.0003517072764225304
step: 110, loss: 0.0030671616550534964
step: 120, loss: 0.003373590996488929
step: 130, loss: 0.001507626730017364
step: 140, loss: 0.0017059581587091088
step: 150, loss: 0.002581539563834667
step: 160, loss: 0.0026545021682977676
step: 170, loss: 0.0003227461129426956
step: 180, loss: 0.00041636134847067297
step: 190, loss: 0.00021811276383232325
step: 200, loss: 0.00017003931861836463
step: 210, loss: 0.0001854388538049534
step: 220, loss: 0.00014713969721924514
step: 230, loss: 0.00010983741231029853
step: 240, loss: 0.0004935802426189184
step: 250, loss: 0.0007453553844243288
step: 260, loss: 0.00026971701299771667
step: 270, loss: 0.00013913991278968751
step: 280, loss: 0.00024203071370720863
step: 290, loss: 0.000492619292344898
step: 300, loss: 0.005990531295537949
step: 310, loss: 0.000265474256593734
step: 320, loss: 0.0015144102508202195
step: 330, loss: 0.00920630432665348

```

```

step: 340, loss: 0.003893478075042367
step: 350, loss: 0.05724504590034485
step: 360, loss: 0.00015739309310447425
step: 370, loss: 0.001331867417320609
step: 380, loss: 0.0781644657254219
step: 390, loss: 0.031257349997758865
step: 400, loss: 0.005406765732914209
step: 410, loss: 0.0009936001151800156
step: 420, loss: 0.01477260421961546
step: 430, loss: 0.00041938092908822
step: 440, loss: 0.0033781114034354687
step: 450, loss: 0.025791659951210022
step: 460, loss: 0.013790484517812729
step: 470, loss: 0.0011933621717616916
step: 480, loss: 0.0038757133297622204
step: 490, loss: 9.186819806927815e-05
step: 500, loss: 0.005740032065659761
step: 510, loss: 0.00019469964900054038
step: 520, loss: 0.0032051843591034412
step: 530, loss: 0.00134774181060493
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9236533957845432, f1=0.9135109864422627,
best_f1=0.9221501390176089
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002471144776791334
step: 10, loss: 0.020163334906101227
step: 20, loss: 0.0007370056700892746
step: 30, loss: 0.0050429003313183784
step: 40, loss: 0.00018257551710121334
step: 50, loss: 0.0005923824501223862
step: 60, loss: 0.06493059545755386
step: 70, loss: 0.001592601416632533
step: 80, loss: 0.0172104649245739
step: 90, loss: 0.0002041857223957777
step: 100, loss: 0.000512637896463275
step: 110, loss: 0.003528888802975416
step: 120, loss: 0.00039495524833910167
step: 130, loss: 0.0001019253977574408
step: 140, loss: 0.002479284768924117
step: 150, loss: 8.032393088797107e-05
step: 160, loss: 0.0012395933736115694
step: 170, loss: 0.009382322430610657
step: 180, loss: 0.00032843954977579415
step: 190, loss: 0.00013523017696570605

```

```

step: 200, loss: 0.00013968384882900864
step: 210, loss: 0.04833206906914711
step: 220, loss: 0.0005000808741897345
step: 230, loss: 0.003210505237802863
step: 240, loss: 0.003214420983567834
step: 250, loss: 0.0012044652830809355
step: 260, loss: 0.0024406674783676863
step: 270, loss: 5.152962330612354e-05
step: 280, loss: 0.0005291688721626997
step: 290, loss: 0.0010345045011490583
step: 300, loss: 0.0009447406628169119
step: 310, loss: 0.01106424629688263
step: 320, loss: 0.0075982240960001945
step: 330, loss: 6.52777380309999e-05
step: 340, loss: 0.03928464278578758
step: 350, loss: 0.0003111763799097389
step: 360, loss: 0.0031377049162983894
step: 370, loss: 0.0003085888456553221
step: 380, loss: 0.012696553021669388
step: 390, loss: 8.716679440112785e-05
step: 400, loss: 5.697568121831864e-05
step: 410, loss: 0.000663396087475121
step: 420, loss: 6.415953248506412e-05
step: 430, loss: 6.265009142225608e-05
step: 440, loss: 0.00032528553856536746
step: 450, loss: 0.003851813031360507
step: 460, loss: 0.0005932087660767138
step: 470, loss: 0.00452447822317481
step: 480, loss: 0.0008725995430722833
step: 490, loss: 8.286444062832743e-05
step: 500, loss: 6.832183862570673e-05
step: 510, loss: 0.0007246982422657311
step: 520, loss: 0.00021344424749258906
step: 530, loss: 6.364939326886088e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9246417013407305, f1=0.9178338001867413,
best_f1=0.9221501390176089
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007881514611653984
step: 10, loss: 0.03920689597725868
step: 20, loss: 0.1823631376028061
step: 30, loss: 0.0005269771791063249
step: 40, loss: 0.009763183072209358
step: 50, loss: 0.005597787909209728

```

step: 60, loss: 9.633076842874289e-05  
step: 70, loss: 0.0023224372416734695  
step: 80, loss: 9.912950190482661e-05  
step: 90, loss: 0.00020008509454783052  
step: 100, loss: 0.0005941443960182369  
step: 110, loss: 0.0010287371696904302  
step: 120, loss: 8.958532998804003e-05  
step: 130, loss: 0.0005817369092255831  
step: 140, loss: 4.351696770754643e-05  
step: 150, loss: 0.01294347271323204  
step: 160, loss: 0.0002243125782115385  
step: 170, loss: 5.9348836657591164e-05  
step: 180, loss: 0.0001256967370864004  
step: 190, loss: 7.190728501882404e-05  
step: 200, loss: 0.00024748663417994976  
step: 210, loss: 0.0002850741730071604  
step: 220, loss: 0.015471125021576881  
step: 230, loss: 0.0034220970701426268  
step: 240, loss: 0.00021178770111873746  
step: 250, loss: 0.00012308369332458824  
step: 260, loss: 0.051909349858760834  
step: 270, loss: 7.384075433947146e-05  
step: 280, loss: 0.00017660603043623269  
step: 290, loss: 0.00014315736189018935  
step: 300, loss: 7.736100815236568e-05  
step: 310, loss: 0.0684908851981163  
step: 320, loss: 0.01760982722043991  
step: 330, loss: 0.027504952624440193  
step: 340, loss: 0.0011099554831162095  
step: 350, loss: 0.00015969194646459073  
step: 360, loss: 0.000151369022205472  
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0  
step: 370, loss: 0.0002188714424846694  
step: 380, loss: 0.0004778294241987169  
step: 390, loss: 0.00011704413191182539  
step: 400, loss: 0.009822300635278225  
step: 410, loss: 0.07723163068294525  
step: 420, loss: 0.00040198396891355515  
step: 430, loss: 0.002159716095775366  
step: 440, loss: 0.0038711277302354574  
step: 450, loss: 0.0008353673620149493  
step: 460, loss: 0.001487180357798934  
step: 470, loss: 0.0002227729419246316  
step: 480, loss: 0.00012045180483255535  
step: 490, loss: 5.161395529285073e-05  
step: 500, loss: 0.000446044112322852  
step: 510, loss: 0.00014720766921527684  
step: 520, loss: 0.044086527079343796

```
step: 530, loss: 0.00034363463055342436
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9253731343283583, f1=0.9150326797385621,
best_f1=0.9221501390176089
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0027219587936997414
step: 10, loss: 0.0004430196131579578
step: 20, loss: 0.0012885481119155884
step: 30, loss: 0.0006229917053133249
step: 40, loss: 0.007014884147793055
step: 50, loss: 0.00023298617452383041
step: 60, loss: 0.00011924154387088493
step: 70, loss: 0.11960308253765106
step: 80, loss: 0.0002143472374882549
step: 90, loss: 0.00011716580775100738
step: 100, loss: 0.0002158137212973088
step: 110, loss: 0.011531892232596874
step: 120, loss: 0.00021018594270572066
step: 130, loss: 0.00014479851233772933
step: 140, loss: 0.00028913424466736615
step: 150, loss: 0.0008595801191404462
step: 160, loss: 0.0004947487614117563
step: 170, loss: 5.116768443258479e-05
step: 180, loss: 0.0004071613075211644
step: 190, loss: 0.0007497409824281931
step: 200, loss: 7.995540363481268e-05
step: 210, loss: 0.00014581059804186225
step: 220, loss: 0.00015829602489247918
step: 230, loss: 0.0004738368443213403
step: 240, loss: 0.0005971973296254873
step: 250, loss: 0.00030783715192228556
step: 260, loss: 0.0005664763739332557
step: 270, loss: 0.009319024160504341
step: 280, loss: 0.00010213940549874678
step: 290, loss: 0.00013167406723368913
step: 300, loss: 0.0001666475145611912
step: 310, loss: 0.03546954318881035
step: 320, loss: 7.016264135017991e-05
step: 330, loss: 0.0062522077932953835
step: 340, loss: 0.00040550727862864733
step: 350, loss: 9.816965030040592e-05
step: 360, loss: 0.0016687382012605667
step: 370, loss: 6.445313192671165e-05
step: 380, loss: 5.217138823354617e-05
```

```

step: 390, loss: 0.0001309466897509992
step: 400, loss: 0.022099580615758896
step: 410, loss: 0.009297364391386509
step: 420, loss: 0.00047684385208413005
step: 430, loss: 0.00015919168072286993
step: 440, loss: 0.00015740467642899603
step: 450, loss: 0.00013200558896642178
step: 460, loss: 0.00020954178762622178
step: 470, loss: 0.000130602769786492
step: 480, loss: 0.006176129449158907
step: 490, loss: 0.00012182763748569414
step: 500, loss: 0.0022712470963597298
step: 510, loss: 0.047146931290626526
step: 520, loss: 4.524430914898403e-05
step: 530, loss: 0.00021876092068850994
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9216417910447762, f1=0.9197559831065226,
best_f1=0.9221501390176089
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004473845474421978
step: 10, loss: 2.6899660952039994e-05
step: 20, loss: 0.0002884026325773448
step: 30, loss: 4.218130197841674e-05
step: 40, loss: 9.645571844885126e-05
step: 50, loss: 0.00011936404916923493
step: 60, loss: 0.14021827280521393
step: 70, loss: 0.0003186841495335102
step: 80, loss: 0.00184870813973248
step: 90, loss: 0.0006933597032912076
step: 100, loss: 0.0007012906717136502
step: 110, loss: 0.0011910991743206978
step: 120, loss: 0.004851804580539465
step: 130, loss: 0.0001967713178601116
step: 140, loss: 0.0031029116362333298
step: 150, loss: 0.001178399776108563
step: 160, loss: 0.00010265951277688146
step: 170, loss: 4.8250578402075917e-05
step: 180, loss: 0.001341973664239049
step: 190, loss: 0.0001739812141750008
step: 200, loss: 0.003427992109209299
step: 210, loss: 0.0009457440464757383
step: 220, loss: 0.001116727595217526
step: 230, loss: 0.00017144436424132437
step: 240, loss: 0.0025464773643761873

```



```

step: 250, loss: 6.336254591587931e-05
step: 260, loss: 0.009098812937736511
step: 270, loss: 3.368467150721699e-05
step: 280, loss: 0.0002558128326199949
step: 290, loss: 9.031668014358729e-05
step: 300, loss: 0.00010008696699514985
step: 310, loss: 0.00015244940004777163
step: 320, loss: 0.0003669113211799413
step: 330, loss: 0.0003776699595618993
step: 340, loss: 0.00012835857341997325
step: 350, loss: 5.280134791973978e-05
step: 360, loss: 0.0005456664366647601
step: 370, loss: 0.0003302486438769847
step: 380, loss: 8.994395466288552e-05
step: 390, loss: 5.342585427570157e-05
step: 400, loss: 0.0001638864487176761
step: 410, loss: 6.036251215846278e-05
step: 420, loss: 0.00038467353442683816
step: 430, loss: 0.0003461592714302242
step: 440, loss: 4.788575097336434e-05
step: 450, loss: 0.00405902648344636
step: 460, loss: 0.00037256337236613035
step: 470, loss: 3.1045652576722205e-05
step: 480, loss: 0.0004332372627686709
step: 490, loss: 0.001220714533701539
step: 500, loss: 0.0634952262043953
step: 510, loss: 5.410454105003737e-05
step: 520, loss: 3.815232776105404e-05
step: 530, loss: 5.637911817757413e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9242916860195076, f1=0.9216504404265183,
best_f1=0.9221501390176089
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02422056533396244
step: 10, loss: 0.0005937323439866304
step: 20, loss: 2.542041147535201e-05
step: 30, loss: 0.000505044765304774
step: 40, loss: 7.738375279586762e-05
step: 50, loss: 4.237622124492191e-05
step: 60, loss: 7.097741035977378e-05
step: 70, loss: 0.00013894909352529794
step: 80, loss: 0.00012220744974911213
step: 90, loss: 4.300803266232833e-05
step: 100, loss: 0.0007495930185541511

```

```

step: 110, loss: 0.0026095539797097445
step: 120, loss: 0.0016682642744854093
step: 130, loss: 2.8549831768032163e-05
step: 140, loss: 4.7003206418594345e-05
step: 150, loss: 4.787354919244535e-05
step: 160, loss: 0.0001374531420879066
step: 170, loss: 9.251971641788259e-05
step: 180, loss: 2.9000562790315598e-05
step: 190, loss: 0.0008793927845545113
step: 200, loss: 0.00408606743440032
step: 210, loss: 0.027918875217437744
step: 220, loss: 0.0023025651462376118
step: 230, loss: 0.0020874023903161287
step: 240, loss: 2.5975723474402912e-05
step: 250, loss: 0.0005560146528296173
step: 260, loss: 7.781691965647042e-05
step: 270, loss: 0.0005466400762088597
step: 280, loss: 5.6644927099114284e-05
step: 290, loss: 0.0005523339495994151
step: 300, loss: 0.1859869360923767
step: 310, loss: 0.003699993947520852
step: 320, loss: 0.0011167326010763645
step: 330, loss: 0.0016122282249853015
step: 340, loss: 2.827762727974914e-05
step: 350, loss: 8.619282743893564e-05
step: 360, loss: 7.694712985539809e-05
step: 370, loss: 3.20582912536338e-05
step: 380, loss: 0.0001507610286353156
step: 390, loss: 0.0017432249151170254
step: 400, loss: 3.519295569276437e-05
step: 410, loss: 2.0641442461055703e-05
step: 420, loss: 0.06011565774679184
step: 430, loss: 0.013780458830296993
step: 440, loss: 0.00014490728790406138
step: 450, loss: 0.00022575682669412345
step: 460, loss: 0.0011007471475750208
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 470, loss: 0.006845323834568262
step: 480, loss: 0.0004269512137398124
step: 490, loss: 0.017309360206127167
step: 500, loss: 3.7381490983534604e-05
step: 510, loss: 3.2787993404781446e-05
step: 520, loss: 2.066003071377054e-05
step: 530, loss: 9.015079558594152e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.923005132991134, f1=0.9147940074906367,

```

```
best_f1=0.9221501390176089
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015708569844719023
step: 10, loss: 3.6729896237375215e-05
step: 20, loss: 3.071140599786304e-05
step: 30, loss: 3.499656304484233e-05
step: 40, loss: 0.00032170716440305114
step: 50, loss: 3.4218410291941836e-05
step: 60, loss: 3.064681732212193e-05
step: 70, loss: 0.00014095559890847653
step: 80, loss: 0.0013906770618632436
step: 90, loss: 5.642314863507636e-05
step: 100, loss: 2.6158117179875262e-05
step: 110, loss: 2.9199241907917894e-05
step: 120, loss: 0.00011459834786364809
step: 130, loss: 6.255190237425268e-05
step: 140, loss: 3.6945039028069004e-05
step: 150, loss: 0.00038154455251060426
step: 160, loss: 2.9097189326421358e-05
step: 170, loss: 0.00012258913193363696
step: 180, loss: 9.572796989232302e-05
step: 190, loss: 2.554646744101774e-05
step: 200, loss: 0.002606053603813052
step: 210, loss: 0.0002084880106849596
step: 220, loss: 2.2008573068887927e-05
step: 230, loss: 0.0007071886793710291
step: 240, loss: 0.0011810249416157603
step: 250, loss: 4.9959086027229205e-05
step: 260, loss: 3.406108226045035e-05
step: 270, loss: 4.919603088637814e-05
step: 280, loss: 5.095949745737016e-05
step: 290, loss: 0.0007507372065447271
step: 300, loss: 8.326887473231182e-05
step: 310, loss: 3.878883217112161e-05
step: 320, loss: 2.3230439182952978e-05
step: 330, loss: 3.2146632293006405e-05
step: 340, loss: 9.55758587224409e-05
step: 350, loss: 0.00015179775073193014
step: 360, loss: 0.0002164569013984874
step: 370, loss: 2.7290283469483256e-05
step: 380, loss: 6.621591455768794e-05
step: 390, loss: 4.867761526838876e-05
step: 400, loss: 2.0730765754706226e-05
step: 410, loss: 0.0005091385683044791
step: 420, loss: 0.00022110631107352674
step: 430, loss: 0.046089403331279755
```

```

step: 440, loss: 0.000508311262819916
step: 450, loss: 0.0001173553682747297
step: 460, loss: 0.00295285414904356
step: 470, loss: 0.0001569468149682507
step: 480, loss: 0.0004341360763646662
step: 490, loss: 2.0861094526480883e-05
step: 500, loss: 0.00304492749273777
step: 510, loss: 0.0001399241155013442
step: 520, loss: 0.0005627829814329743
step: 530, loss: 0.0009965025819838047
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9183766529867762, f1=0.9177011494252874,
best_f1=0.9221501390176089
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001433432917110622
step: 10, loss: 1.8261071090819314e-05
step: 20, loss: 0.00014201873273123056
step: 30, loss: 0.0004075320321135223
step: 40, loss: 3.381259011803195e-05
step: 50, loss: 0.00021254798048175871
step: 60, loss: 5.6644552387297153e-05
step: 70, loss: 2.343134838156402e-05
step: 80, loss: 0.03272073343396187
step: 90, loss: 0.001522985752671957
step: 100, loss: 2.450366628181655e-05
step: 110, loss: 2.0403040252858773e-05
step: 120, loss: 0.00012936948041897267
step: 130, loss: 0.0013003553031012416
step: 140, loss: 4.011933197034523e-05
step: 150, loss: 3.534613279043697e-05
step: 160, loss: 1.9486677047098055e-05
step: 170, loss: 7.695468957535923e-05
step: 180, loss: 6.815678352722898e-05
step: 190, loss: 0.00024802895495668054
step: 200, loss: 0.0002232338592875749
step: 210, loss: 0.0006556735606864095
step: 220, loss: 2.9334349164855666e-05
step: 230, loss: 4.525238909991458e-05
step: 240, loss: 5.465609865495935e-05
step: 250, loss: 0.040094465017318726
step: 260, loss: 2.108826083713211e-05
step: 270, loss: 5.506797970156185e-05
step: 280, loss: 1.7013044271152467e-05
step: 290, loss: 5.5376563977915794e-05

```

```

step: 300, loss: 6.410448986571282e-05
step: 310, loss: 2.2049309336580336e-05
step: 320, loss: 3.070240200031549e-05
step: 330, loss: 0.00046100173494778574
step: 340, loss: 3.4786546166287735e-05
step: 350, loss: 0.004847644828259945
step: 360, loss: 2.7487014449434355e-05
step: 370, loss: 4.8161327868001536e-05
step: 380, loss: 0.0018186670495197177
step: 390, loss: 3.3949942007893696e-05
step: 400, loss: 3.24626817018725e-05
step: 410, loss: 4.582193651003763e-05
step: 420, loss: 7.384148921119049e-05
step: 430, loss: 0.0003421305736992508
step: 440, loss: 3.016217306139879e-05
step: 450, loss: 0.0002539507986512035
step: 460, loss: 1.8760254533845e-05
step: 470, loss: 0.00012038640852551907
step: 480, loss: 0.00028088325052522123
step: 490, loss: 0.00027587925433181226
step: 500, loss: 0.00011949988402193412
step: 510, loss: 2.311862954229582e-05
step: 520, loss: 2.693279202503618e-05
step: 530, loss: 3.4194519685115665e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9229357798165138, f1=0.92201199815413,
best_f1=0.9221501390176089
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00031624347320757806
step: 10, loss: 3.3502921723993495e-05
step: 20, loss: 2.0801368009415455e-05
step: 30, loss: 7.079974602675065e-05
step: 40, loss: 0.00027928315103054047
step: 50, loss: 4.765972335007973e-05
step: 60, loss: 4.172782428213395e-05
step: 70, loss: 3.8392328860936686e-05
step: 80, loss: 2.4515558834536932e-05
step: 90, loss: 1.7426642443751916e-05
step: 100, loss: 3.996980740339495e-05
step: 110, loss: 2.0171934011159465e-05
step: 120, loss: 1.7690743334242143e-05
step: 130, loss: 3.420640132389963e-05
step: 140, loss: 5.305768718244508e-05
step: 150, loss: 0.00015007522597443312

```

```

step: 160, loss: 0.0003516919387038797
step: 170, loss: 0.0005092715146020055
step: 180, loss: 2.827384196280036e-05
step: 190, loss: 7.965961412992328e-05
step: 200, loss: 0.0006097860168665648
step: 210, loss: 0.00014745121006853878
step: 220, loss: 3.6131164961261675e-05
step: 230, loss: 0.0015166652156040072
step: 240, loss: 1.9710201740963385e-05
step: 250, loss: 3.223667954443954e-05
step: 260, loss: 1.4211831512511708e-05
step: 270, loss: 0.0005494448123499751
step: 280, loss: 1.9594461264205165e-05
step: 290, loss: 2.482815398252569e-05
step: 300, loss: 0.00031481022597290576
step: 310, loss: 2.76509890682064e-05
step: 320, loss: 4.476575122680515e-05
step: 330, loss: 0.0007180629181675613
step: 340, loss: 3.10669893224258e-05
step: 350, loss: 0.022973207756876945
step: 360, loss: 0.0014326077653095126
step: 370, loss: 5.56168015464209e-05
step: 380, loss: 1.943433926498983e-05
step: 390, loss: 1.839517972257454e-05
step: 400, loss: 3.0345205232151784e-05
step: 410, loss: 2.8291460694163106e-05
step: 420, loss: 2.5260482289013453e-05
step: 430, loss: 2.983456761285197e-05
step: 440, loss: 1.622710442461539e-05
step: 450, loss: 2.829109325830359e-05
step: 460, loss: 4.3922111217398196e-05
step: 470, loss: 1.997457911784295e-05
step: 480, loss: 2.5069808543776162e-05
step: 490, loss: 1.6223413695115596e-05
step: 500, loss: 3.6807025026064366e-05
step: 510, loss: 3.652687883004546e-05
step: 520, loss: 3.597645627451129e-05
step: 530, loss: 2.0820329154958017e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9208699676075892, f1=0.9202797202797203,
best_f1=0.9221501390176089
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.8007700418820605e-05
step: 10, loss: 0.00012616149615496397

```

step: 20, loss: 2.8519860279629938e-05  
step: 30, loss: 0.00028670381288975477  
step: 40, loss: 1.3664224752574228e-05  
step: 50, loss: 0.00023239466827362776  
step: 60, loss: 4.9195397878065705e-05  
step: 70, loss: 0.0009679876966401935  
step: 80, loss: 2.5178049327223562e-05  
step: 90, loss: 1.9020832041860558e-05  
step: 100, loss: 1.5456009350600652e-05  
step: 110, loss: 5.591521767200902e-05  
step: 120, loss: 0.0017573690274730325  
step: 130, loss: 1.5675785107305273e-05  
step: 140, loss: 2.1807023586006835e-05  
step: 150, loss: 0.00012442399747669697  
step: 160, loss: 0.00011350367276463658  
step: 170, loss: 1.90693826880306e-05  
step: 180, loss: 1.950488513102755e-05  
step: 190, loss: 3.9802573155611753e-05  
step: 200, loss: 3.5844193917000666e-05  
step: 210, loss: 5.4709882533643395e-05  
step: 220, loss: 1.3582207429863047e-05  
step: 230, loss: 1.861117016233038e-05  
step: 240, loss: 2.6649835490388796e-05  
step: 250, loss: 0.00014735541481059045  
step: 260, loss: 6.290794408414513e-05  
step: 270, loss: 3.730268872459419e-05  
step: 280, loss: 9.609981498215348e-05  
step: 290, loss: 0.00141845119651407  
step: 300, loss: 3.353522333782166e-05  
step: 310, loss: 7.655635999981314e-05  
step: 320, loss: 6.888830102980137e-05  
step: 330, loss: 4.078457277501002e-05  
step: 340, loss: 0.00047184014692902565  
step: 350, loss: 3.887376806233078e-05  
step: 360, loss: 4.932360025122762e-05  
step: 370, loss: 1.704667738522403e-05  
step: 380, loss: 1.7028041838784702e-05  
step: 390, loss: 1.6972177036223002e-05  
step: 400, loss: 1.7620281141716987e-05  
step: 410, loss: 4.2350118746981025e-05  
step: 420, loss: 0.00010865418880712241  
step: 430, loss: 1.9069440895691514e-05  
step: 440, loss: 3.5335691791260615e-05  
step: 450, loss: 0.00043324960279278457  
step: 460, loss: 1.3597164979728404e-05  
step: 470, loss: 0.00011758563050534576  
step: 480, loss: 0.0012760082026943564  
step: 490, loss: 2.3558219254482538e-05

```

step: 500, loss: 3.059715891140513e-05
step: 510, loss: 2.26156316784909e-05
step: 520, loss: 1.9069409972871654e-05
step: 530, loss: 1.97772060346324e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9223659889094269, f1=0.9190697674418604,
best_f1=0.9221501390176089
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:15, 371.27it/s]
load_f1 = 0.9260808926080892
real_f1 = 0.9255121042830541
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:11, 393.33it/s]

```



## 4 DITTO

### 4.1 DITTO STRUCTURED

#### 4.1.1 Beer - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/Beer \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm distilbert \
  --fp16 \
  --save_model \
  --dk general \
  --dk product \
  --summarize \
  --da entry_swap \
  --da attr_del

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Beer.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/Beer \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm distilbert \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/
```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

|                 |                 |
|-----------------|-----------------|
| enabled         | : True          |
| opt_level       | : 02            |
| cast_model_type | : torch.float16 |

```

patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled               : True
opt_level             : 02
cast_model_type       : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale            : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.8596862554550171
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.288659793814433, f1=0.33766233766233766,
best_f1=0.33766233766233766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3933474123477936
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

```

UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.3333333333333333, f1=0.0, best_f1=0.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3005567789077759
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.4057971014492754, f1=0.25641025641025644,
best_f1=0.25641025641025644
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.32269027829170227
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.4, f1=0.16666666666666666, best_f1=0.25641025641025644
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23227167129516602
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.3943661971830986, f1=0.20512820512820512,
best_f1=0.25641025641025644
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24663251638412476
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 6: dev_f1=0.4444444444444444, f1=0.19047619047619047,
best_f1=0.19047619047619047
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1841350495815277
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.43137254901960786, f1=0.22222222222222224,
best_f1=0.19047619047619047
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3036705553531647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 8: dev_f1=0.46666666666666667, f1=0.0, best_f1=0.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16513893008232117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.5142857142857143, f1=0.21052631578947364,
best_f1=0.21052631578947364
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23579050600528717
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 10: dev_f1=0.5294117647058824, f1=0.21052631578947364,
best_f1=0.21052631578947364
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2183135449886322
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 11: dev_f1=0.5454545454545454, f1=0.3, best_f1=0.3
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19940254092216492
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.5306122448979592, f1=0.3870967741935484, best_f1=0.3
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13183291256427765
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 13: dev_f1=0.5555555555555555, f1=0.41666666666666663,
best_f1=0.41666666666666663
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19960381090641022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.5405405405405405, f1=0.4799999999999999,
best_f1=0.41666666666666663
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2713104486465454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.5405405405405405, f1=0.4799999999999999,

```

```

best_f1=0.41666666666666663
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
91it [00:00, 137394.41it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
load_f1 = 0.5405405405405405
real_f1 = 0.5581395348837208
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:10, 417.63it/s]

```

#### 4.1.2 DBLP-ACM - Running the matcher

```

[ ]: # Step 1:
    # train the wdc model with the save_model flag on
    # it will produce two files *_dev.pt and *_test.pt which are the
    # best checkpoints on the dev set and the test set.
    !CUDA_VISIBLE_DEVICES=0 python train_ditto.py \

```

```

--task Structured/DBLP-ACM \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm distilbert \
--fp16 \
--save_model \
--dk general \
--dk product \
--summarize \
--da entry_swap \
--da attr_del

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-ACM.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/DBLP-ACM \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm distilbert \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True

```

```

master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.8020479679107666
step: 10, loss: 0.4745992422103882
step: 20, loss: 0.5912355184555054
step: 30, loss: 0.42239007353782654
step: 40, loss: 0.18217749893665314
step: 50, loss: 0.22298043966293335
step: 60, loss: 0.12384632974863052
step: 70, loss: 0.1696169674396515
step: 80, loss: 0.21362102031707764
step: 90, loss: 0.01239192858338356
step: 100, loss: 0.08137055486440659
step: 110, loss: 0.1059483215212822
step: 120, loss: 0.011148100718855858
step: 130, loss: 0.05539269372820854
step: 140, loss: 0.06030699238181114
step: 150, loss: 0.01093105599284172
step: 160, loss: 0.26371586322784424
step: 170, loss: 0.010131694376468658
step: 180, loss: 0.007465628907084465
step: 190, loss: 0.06464818120002747
step: 200, loss: 0.00173112319316715
step: 210, loss: 0.048650480806827545
step: 220, loss: 0.006729885004460812
step: 230, loss: 0.02425318956375122
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.

```



```

    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9751693002257337, f1=0.9714285714285715,
best_f1=0.9714285714285715
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05784335732460022
step: 10, loss: 0.003061986295506358
step: 20, loss: 0.030812660232186317
step: 30, loss: 0.0359148345887661
step: 40, loss: 0.006476760841906071
step: 50, loss: 0.006236114539206028
step: 60, loss: 0.0024436581879854202
step: 70, loss: 0.08714134991168976
step: 80, loss: 0.010633844882249832
step: 90, loss: 0.007826963439583778
step: 100, loss: 0.010382694192230701
step: 110, loss: 0.034678634256124496
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 120, loss: 0.016301453113555908
step: 130, loss: 0.008114341646432877
step: 140, loss: 0.14540046453475952
step: 150, loss: 0.004081345163285732
step: 160, loss: 0.09999271482229233
step: 170, loss: 0.07239069044589996
step: 180, loss: 0.0017072852933779359
step: 190, loss: 0.05883199721574783
step: 200, loss: 0.0018053697422146797
step: 210, loss: 0.04567509889602661
step: 220, loss: 0.001388300908729434
step: 230, loss: 0.0541815310716629
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9842342342342343, f1=0.9730337078651685,
best_f1=0.9730337078651685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15183447301387787
step: 10, loss: 0.019088968634605408
step: 20, loss: 0.010367762297391891
step: 30, loss: 0.0751875787973404
step: 40, loss: 0.006573644001036882

```

```

step: 50, loss: 0.027359219267964363
step: 60, loss: 0.0018400249537080526
step: 70, loss: 0.003817599965259433
step: 80, loss: 0.041870348155498505
step: 90, loss: 0.01193295232951641
step: 100, loss: 0.0008268853998742998
step: 110, loss: 0.014740916900336742
step: 120, loss: 0.011697480455040932
step: 130, loss: 0.0004491850268095732
step: 140, loss: 0.0021991345565766096
step: 150, loss: 0.0019960827194154263
step: 160, loss: 0.0030724816024303436
step: 170, loss: 0.02818216197192669
step: 180, loss: 0.0011472590267658234
step: 190, loss: 0.00862235575914383
step: 200, loss: 0.0016779988072812557
step: 210, loss: 0.0034288305323570967
step: 220, loss: 0.020024845376610756
step: 230, loss: 0.0660252645611763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.9807474518686297, f1=0.9738933030646991,
best_f1=0.9730337078651685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0031823047902435064
step: 10, loss: 0.0008772385772317648
step: 20, loss: 0.014154452830553055
step: 30, loss: 0.0016602774849161506
step: 40, loss: 0.0022204702254384756
step: 50, loss: 0.005000504665076733
step: 60, loss: 0.0006700681988149881
step: 70, loss: 0.001312808133661747
step: 80, loss: 0.03290534391999245
step: 90, loss: 0.0007959209033288062
step: 100, loss: 0.0016780032310634851
step: 110, loss: 0.001929575577378273
step: 120, loss: 0.006014447659254074
step: 130, loss: 0.0011254489654675126
step: 140, loss: 0.02098388597369194
step: 150, loss: 0.007401828654110432
step: 160, loss: 0.004882253240793943
step: 170, loss: 0.05039415881037712
step: 180, loss: 0.054211001843214035
step: 190, loss: 0.0024052667431533337
step: 200, loss: 0.0010526629630476236

```

```

step: 210, loss: 0.00035547022707760334
step: 220, loss: 0.0007753671961836517
step: 230, loss: 0.0006131173577159643
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.980963045912654, f1=0.9776286353467561,
best_f1=0.9730337078651685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021798153466079384
step: 10, loss: 0.0003465202171355486
step: 20, loss: 0.041359033435583115
step: 30, loss: 0.0027997829020023346
step: 40, loss: 0.002064955187961459
step: 50, loss: 0.001480601029470563
step: 60, loss: 0.125356987118721
step: 70, loss: 0.0023274749983102083
step: 80, loss: 0.0016871653497219086
step: 90, loss: 0.052583593875169754
step: 100, loss: 0.06472323089838028
step: 110, loss: 0.07661908119916916
step: 120, loss: 0.009556103497743607
step: 130, loss: 0.0005462564877234399
step: 140, loss: 0.00032988403108902276
step: 150, loss: 0.000626266177278012
step: 160, loss: 0.006265060044825077
step: 170, loss: 0.031684596091508865
step: 180, loss: 0.00032641980214975774
step: 190, loss: 0.0031627153512090445
step: 200, loss: 0.00949519407004118
step: 210, loss: 0.000668322725687176
step: 220, loss: 0.002232800703495741
step: 230, loss: 0.00031548659899272025
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.983050847457627, f1=0.9784335981838819,
best_f1=0.9730337078651685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0020327819511294365
step: 10, loss: 0.00021843163995072246
step: 20, loss: 0.0002693989663384855
step: 30, loss: 0.002372537972405553
step: 40, loss: 0.0009594102157279849

```

```

step: 50, loss: 0.005122047383338213
step: 60, loss: 0.009305641055107117
step: 70, loss: 0.014688120223581791
step: 80, loss: 0.00020275839779060334
step: 90, loss: 0.00012801079719793051
step: 100, loss: 0.00013026634405832738
step: 110, loss: 0.119882732629776
step: 120, loss: 0.00027306799893267453
step: 130, loss: 0.0165094081312418
step: 140, loss: 0.1767866462469101
step: 150, loss: 0.011009926907718182
step: 160, loss: 0.021645771339535713
step: 170, loss: 0.0007107619312591851
step: 180, loss: 0.0015009189955890179
step: 190, loss: 0.005233149044215679
step: 200, loss: 0.006273314356803894
step: 210, loss: 0.0014306209050118923
step: 220, loss: 0.058156244456768036
step: 230, loss: 0.015306518413126469
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9743016759776536, f1=0.9752808988764046,
best_f1=0.9730337078651685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12093686312437057
step: 10, loss: 0.0037871277891099453
step: 20, loss: 0.0009525156929157674
step: 30, loss: 0.00018931693921331316
step: 40, loss: 0.012622030451893806
step: 50, loss: 0.0001457048492738977
step: 60, loss: 0.030323363840579987
step: 70, loss: 0.0005133946542628109
step: 80, loss: 7.549483416369185e-05
step: 90, loss: 0.00022051935957279056
step: 100, loss: 0.0003113399143330753
step: 110, loss: 0.0014404846588149667
step: 120, loss: 0.006114591844379902
step: 130, loss: 0.0007457525935024023
step: 140, loss: 0.00043282273691147566
step: 150, loss: 0.0002488266909494996
step: 160, loss: 0.00010669963376130909
step: 170, loss: 9.403886360814795e-05
step: 180, loss: 0.00026356489979662
step: 190, loss: 0.0004963133833371103
step: 200, loss: 0.021527640521526337

```

```

step: 210, loss: 0.0007176004583016038
step: 220, loss: 0.0005450207972899079
step: 230, loss: 0.015237698331475258
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9808342728297633, f1=0.976324689966178,
best_f1=0.9730337078651685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015364564023911953
step: 10, loss: 0.16459636390209198
step: 20, loss: 0.0033070191275328398
step: 30, loss: 0.00036304484819993377
step: 40, loss: 0.00019510347920004278
step: 50, loss: 0.0006928734946995974
step: 60, loss: 0.00031296242377720773
step: 70, loss: 0.0003886502527166158
step: 80, loss: 0.0056176213547587395
step: 90, loss: 0.0002076077798847109
step: 100, loss: 0.000749876257032156
step: 110, loss: 0.00011374877067282796
step: 120, loss: 0.00012022646114928648
step: 130, loss: 0.011528126895427704
step: 140, loss: 0.0013412287225946784
step: 150, loss: 0.004214572254568338
step: 160, loss: 0.005344008095562458
step: 170, loss: 0.0001968354918062687
step: 180, loss: 0.0012112853582948446
step: 190, loss: 0.00032689995714463294
step: 200, loss: 0.00029717496363446116
step: 210, loss: 5.376426997827366e-05
step: 220, loss: 0.08138170838356018
step: 230, loss: 0.022769656032323837
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.966740576496674, f1=0.9753914988814317,
best_f1=0.9730337078651685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08760900050401688
step: 10, loss: 0.00037707117735408247
step: 20, loss: 0.006503593642264605
step: 30, loss: 0.010149745270609856
step: 40, loss: 0.00031222912366501987

```

```

step: 50, loss: 0.0004506414697971195
step: 60, loss: 0.0007703315350227058
step: 70, loss: 0.0017254779813811183
step: 80, loss: 0.10682922601699829
step: 90, loss: 0.011766912415623665
step: 100, loss: 0.00036785713746212423
step: 110, loss: 0.0003553294518496841
step: 120, loss: 0.0229186974465847
step: 130, loss: 0.0001051352737704292
step: 140, loss: 0.0001438283215975389
step: 150, loss: 6.162247154861689e-05
step: 160, loss: 0.00024034458328969777
step: 170, loss: 9.024218161357567e-05
step: 180, loss: 0.001964292488992214
step: 190, loss: 0.00018586573423817754
step: 200, loss: 0.001445620320737362
step: 210, loss: 0.00013947216211818159
step: 220, loss: 0.023847943171858788
step: 230, loss: 0.012635773979127407
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9808773903262092, f1=0.9765363128491621,
best_f1=0.9730337078651685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014579702110495418
step: 10, loss: 0.0001632572675589472
step: 20, loss: 9.796579251997173e-05
step: 30, loss: 0.03720811381936073
step: 40, loss: 0.00042584422044456005
step: 50, loss: 0.028043750673532486
step: 60, loss: 0.024869130924344063
step: 70, loss: 0.0005121593712829053
step: 80, loss: 0.001643452444113791
step: 90, loss: 0.0025004432536661625
step: 100, loss: 0.00026101100957021117
step: 110, loss: 6.695720367133617e-05
step: 120, loss: 0.0031912634149193764
step: 130, loss: 0.0015964745543897152
step: 140, loss: 0.0342390350997448
step: 150, loss: 6.384150037774816e-05
step: 160, loss: 0.0002594390243757516
step: 170, loss: 0.00024308323918376118
step: 180, loss: 0.00010131345334229991
step: 190, loss: 0.0004972996539436281
step: 200, loss: 4.7517416533082724e-05

```

```

step: 210, loss: 0.00014083947462495416
step: 220, loss: 0.00014627024938818067
step: 230, loss: 0.0014925571158528328
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9733924611973392, f1=0.9689578713968958,
best_f1=0.9730337078651685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.241396815516055e-05
step: 10, loss: 0.00013463458162732422
step: 20, loss: 4.526988777797669e-05
step: 30, loss: 0.00010191109322477132
step: 40, loss: 0.00032052365713752806
step: 50, loss: 0.0001491962611908093
step: 60, loss: 0.00012785744911525398
step: 70, loss: 0.0013814510311931372
step: 80, loss: 0.001058558584190905
step: 90, loss: 0.0006716050556860864
step: 100, loss: 8.111798524623737e-05
step: 110, loss: 3.860762808471918e-05
step: 120, loss: 0.05255600064992905
step: 130, loss: 7.282453589141369e-05
step: 140, loss: 8.45932518132031e-05
step: 150, loss: 0.0002702594792935997
step: 160, loss: 0.00020855030743405223
step: 170, loss: 0.009024057537317276
step: 180, loss: 0.00018238899065181613
step: 190, loss: 0.00818968191742897
step: 200, loss: 5.957148096058518e-05
step: 210, loss: 0.00023745444195810705
step: 220, loss: 0.00019548939599189907
step: 230, loss: 0.00021339836530387402
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9784335981838819, f1=0.9739524348810873,
best_f1=0.9730337078651685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003032856620848179
step: 10, loss: 0.0001781786122592166
step: 20, loss: 5.423220864031464e-05
step: 30, loss: 0.0002112796064466238
step: 40, loss: 0.00028058799216523767

```

```

step: 50, loss: 4.3736392399296165e-05
step: 60, loss: 0.0003433093661442399
step: 70, loss: 0.0006263102986849844
step: 80, loss: 8.257262379629537e-05
step: 90, loss: 3.74567971448414e-05
step: 100, loss: 6.804405711591244e-05
step: 110, loss: 0.00011303090286673978
step: 120, loss: 0.00014370724966283888
step: 130, loss: 7.919850031612441e-05
step: 140, loss: 5.18093365826644e-05
step: 150, loss: 9.603767830412835e-05
step: 160, loss: 0.023675521835684776
step: 170, loss: 0.00015160006296355277
step: 180, loss: 0.011315375566482544
step: 190, loss: 0.007809133268892765
step: 200, loss: 0.00019031365809496492
step: 210, loss: 6.073430267861113e-05
step: 220, loss: 0.00023701971804257482
step: 230, loss: 6.337684317259118e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9797297297297298, f1=0.9752252252252253,
best_f1=0.9730337078651685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003789758193306625
step: 10, loss: 8.431615424342453e-05
step: 20, loss: 0.00023386176326312125
step: 30, loss: 0.018070271238684654
step: 40, loss: 0.0005563533050008118
step: 50, loss: 0.0001548086729599163
step: 60, loss: 0.001228284789249301
step: 70, loss: 0.0002959298144560307
step: 80, loss: 0.001246076193638146
step: 90, loss: 6.978656892897561e-05
step: 100, loss: 0.00010049752745544538
step: 110, loss: 0.00011486501898616552
step: 120, loss: 7.787361391820014e-05
step: 130, loss: 0.0001617896050447598
step: 140, loss: 0.00012482916645240039
step: 150, loss: 0.029283102601766586
step: 160, loss: 7.345994526986033e-05
step: 170, loss: 5.564924140344374e-05
step: 180, loss: 0.0002858610823750496
step: 190, loss: 0.0014896874781697989
step: 200, loss: 9.302866237703711e-05

```



```

step: 210, loss: 5.367171615944244e-05
step: 220, loss: 7.967640704009682e-05
step: 230, loss: 0.00015569329843856394
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9785310734463276, f1=0.9807909604519773,
best_f1=0.9730337078651685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.505141623667441e-05
step: 10, loss: 3.0520648579113185e-05
step: 20, loss: 0.026968609541654587
step: 30, loss: 5.4140433348948136e-05
step: 40, loss: 3.1619576475350186e-05
step: 50, loss: 0.00011232777615077794
step: 60, loss: 0.00010530583676882088
step: 70, loss: 5.272206908557564e-05
step: 80, loss: 0.00012418739788699895
step: 90, loss: 0.00011667335638776422
step: 100, loss: 0.011289753951132298
step: 110, loss: 0.016817284747958183
step: 120, loss: 0.0003885116020683199
step: 130, loss: 0.00010008481331169605
step: 140, loss: 4.3192387238377705e-05
step: 150, loss: 5.9759357100119814e-05
step: 160, loss: 5.787699410575442e-05
step: 170, loss: 5.860508099431172e-05
step: 180, loss: 7.466526585631073e-05
step: 190, loss: 0.005753824487328529
step: 200, loss: 4.237337634549476e-05
step: 210, loss: 0.0002570290525909513
step: 220, loss: 0.0004227610770612955
step: 230, loss: 9.26023640204221e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9819004524886877, f1=0.976324689966178,
best_f1=0.9730337078651685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.365015774965286e-05
step: 10, loss: 7.177048973971978e-05
step: 20, loss: 5.670678729075007e-05
step: 30, loss: 0.01363381277769804
step: 40, loss: 0.00011761841597035527

```

```

step: 50, loss: 5.9004400100093335e-05
step: 60, loss: 8.209787483792752e-05
step: 70, loss: 6.459909491240978e-05
step: 80, loss: 0.0001331645908067003
step: 90, loss: 3.354166619828902e-05
step: 100, loss: 5.019840318709612e-05
step: 110, loss: 7.704889867454767e-05
step: 120, loss: 8.536609675502405e-05
step: 130, loss: 0.16809290647506714
step: 140, loss: 0.015576244331896305
step: 150, loss: 4.6716875658603385e-05
step: 160, loss: 9.631193825043738e-05
step: 170, loss: 0.0009652271983213723
step: 180, loss: 0.00011930568143725395
step: 190, loss: 0.010446599684655666
step: 200, loss: 0.00023010144650470465
step: 210, loss: 0.029045086354017258
step: 220, loss: 0.00012736130156554282
step: 230, loss: 0.00010683193249860778
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9807037457434733, f1=0.9807474518686297,
best_f1=0.9730337078651685
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type       : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type       : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2473it [00:08, 302.68it/s]
load_f1 = 0.9788182831661093
real_f1 = 0.9821428571428571
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:10, 404.68it/s]

```

### 4.1.3 DBLP-GoogleScholar - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/DBLP-GoogleScholar \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm distilbert \
  --fp16 \
  --save_model \
  --dk general \
  --dk product \
  --summarize \
  --da entry_swap \
  --da attr_del

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-GoogleScholar.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/DBLP-GoogleScholar \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm distilbert \

```

```
--use_gpu \  
--fp16 \  
--checkpoint_path checkpoints/
```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```
enabled          : True  
opt_level        : 02  
cast_model_type  : torch.float16  
patch_torch_functions : False  
keep_batchnorm_fp32 : True  
master_weights   : True  
loss_scale       : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```
enabled          : True  
opt_level        : 02  
cast_model_type  : torch.float16  
patch_torch_functions : False  
keep_batchnorm_fp32 : True  
master_weights   : True  
loss_scale       : dynamic
```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback.

Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning:

An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0

/usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`.

In PyTorch 1.1.0 and later, you should call them in the opposite order:

`optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

"<https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>",  
UserWarning)

step: 0, loss: 0.7900460362434387

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

step: 10, loss: 0.46689584851264954

step: 20, loss: 0.5051938891410828

step: 30, loss: 0.4443317949771881

step: 40, loss: 0.46875447034835815

step: 50, loss: 0.3734215497970581

step: 60, loss: 0.2370021790266037  
step: 70, loss: 0.1509532481431961  
step: 80, loss: 0.21432362496852875  
step: 90, loss: 0.1621088832616806  
step: 100, loss: 0.3517465889453888  
step: 110, loss: 0.11890513449907303  
step: 120, loss: 0.10046876966953278  
step: 130, loss: 0.02857922576367855  
step: 140, loss: 0.22542434930801392  
step: 150, loss: 0.037938524037599564  
step: 160, loss: 0.11769022792577744  
step: 170, loss: 0.2711820900440216  
step: 180, loss: 0.0979599803686142  
step: 190, loss: 0.15175242722034454  
step: 200, loss: 0.08495516330003738  
step: 210, loss: 0.1438191682100296  
step: 220, loss: 0.15967977046966553  
step: 230, loss: 0.213607519865036  
step: 240, loss: 0.12437507510185242  
step: 250, loss: 0.11595332622528076  
step: 260, loss: 0.08560099452733994  
step: 270, loss: 0.03618171066045761  
step: 280, loss: 0.1896144151687622  
step: 290, loss: 0.14688344299793243  
step: 300, loss: 0.2524159550666809  
step: 310, loss: 0.14268968999385834  
step: 320, loss: 0.12520122528076172  
step: 330, loss: 0.1744406670331955  
step: 340, loss: 0.22157365083694458  
step: 350, loss: 0.052638038992881775  
step: 360, loss: 0.14949259161949158  
step: 370, loss: 0.21332328021526337  
step: 380, loss: 0.1341344714164734  
step: 390, loss: 0.05119006335735321  
step: 400, loss: 0.052691515535116196  
step: 410, loss: 0.05549358204007149  
step: 420, loss: 0.043754056096076965  
step: 430, loss: 0.03718699514865875  
step: 440, loss: 0.11878734827041626  
step: 450, loss: 0.05002707988023758  
step: 460, loss: 0.10804297029972076  
step: 470, loss: 0.25801509618759155  
step: 480, loss: 0.2894570529460907  
step: 490, loss: 0.07503914833068848  
step: 500, loss: 0.044110145419836044  
step: 510, loss: 0.09181728959083557  
step: 520, loss: 0.11678465455770493  
step: 530, loss: 0.106493279337883

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9136258660508083, f1=0.908167974157822,
best_f1=0.908167974157822
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17477324604988098
step: 10, loss: 0.13941359519958496
step: 20, loss: 0.08649608492851257
step: 30, loss: 0.05165763944387436
step: 40, loss: 0.009832429699599743
step: 50, loss: 0.138026162981987
step: 60, loss: 0.12031237781047821
step: 70, loss: 0.14213904738426208
step: 80, loss: 0.07526498287916183
step: 90, loss: 0.057491790503263474
step: 100, loss: 0.25381824374198914
step: 110, loss: 0.02874135971069336
step: 120, loss: 0.08227874338626862
step: 130, loss: 0.052860382944345474
step: 140, loss: 0.02500414103269577
step: 150, loss: 0.039350006729364395
step: 160, loss: 0.2054290622472763
step: 170, loss: 0.06220805272459984
step: 180, loss: 0.023666661232709885
step: 190, loss: 0.029895341023802757
step: 200, loss: 0.020866364240646362
step: 210, loss: 0.03131403401494026
step: 220, loss: 0.16287405788898468
step: 230, loss: 0.14700309932231903
step: 240, loss: 0.1098923459649086
step: 250, loss: 0.06801149994134903
step: 260, loss: 0.003441540291532874
step: 270, loss: 0.1075829491019249
step: 280, loss: 0.11227015405893326
step: 290, loss: 0.046844515949487686
step: 300, loss: 0.0257828738540411
step: 310, loss: 0.05213114991784096
step: 320, loss: 0.14859260618686676
step: 330, loss: 0.044948793947696686
step: 340, loss: 0.0295257531106472
step: 350, loss: 0.04936151206493378

```

```

step: 360, loss: 0.13283823430538177
step: 370, loss: 0.02165457420051098
step: 380, loss: 0.06957296282052994
step: 390, loss: 0.0656118392944336
step: 400, loss: 0.049504999071359634
step: 410, loss: 0.0015569019597023726
step: 420, loss: 0.09481428563594818
step: 430, loss: 0.015606123954057693
step: 440, loss: 0.015924254432320595
step: 450, loss: 0.036052118986845016
step: 460, loss: 0.2292410135269165
step: 470, loss: 0.026611916720867157
step: 480, loss: 0.19425393640995026
step: 490, loss: 0.04931741580367088
step: 500, loss: 0.029209818691015244
step: 510, loss: 0.04502727463841438
step: 520, loss: 0.005374900996685028
step: 530, loss: 0.16554328799247742
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9229357798165138, f1=0.9174311926605505,
best_f1=0.9174311926605505
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11508973687887192
step: 10, loss: 0.036946944892406464
step: 20, loss: 0.22490394115447998
step: 30, loss: 0.178684800863266
step: 40, loss: 0.06152134761214256
step: 50, loss: 0.10908330976963043
step: 60, loss: 0.010602934285998344
step: 70, loss: 0.04171386733651161
step: 80, loss: 0.008699973113834858
step: 90, loss: 0.11820036172866821
step: 100, loss: 0.027344712987542152
step: 110, loss: 0.013698202557861805
step: 120, loss: 0.03908466175198555
step: 130, loss: 0.02105667255818844
step: 140, loss: 0.048006340861320496
step: 150, loss: 0.05163617804646492
step: 160, loss: 0.016860509291291237
step: 170, loss: 0.006147331092506647

```

```

step: 180, loss: 0.06887941062450409
step: 190, loss: 0.0027950310613960028
step: 200, loss: 0.007830890826880932
step: 210, loss: 0.024179818108677864
step: 220, loss: 0.05408499389886856
step: 230, loss: 0.019462989643216133
step: 240, loss: 0.0368020236492157
step: 250, loss: 0.03578623756766319
step: 260, loss: 0.029804375022649765
step: 270, loss: 0.0025573261082172394
step: 280, loss: 0.01914878562092781
step: 290, loss: 0.023963388055562973
step: 300, loss: 0.25176605582237244
step: 310, loss: 0.10640329867601395
step: 320, loss: 0.047356992959976196
step: 330, loss: 0.0074360002763569355
step: 340, loss: 0.008377100341022015
step: 350, loss: 0.018160080537199974
step: 360, loss: 0.017844688147306442
step: 370, loss: 0.025363342836499214
step: 380, loss: 0.038211971521377563
step: 390, loss: 0.013088409788906574
step: 400, loss: 0.05797416716814041
step: 410, loss: 0.013538255356252193
step: 420, loss: 0.05690200254321098
step: 430, loss: 0.008335970342159271
step: 440, loss: 0.06954887509346008
step: 450, loss: 0.11972086131572723
step: 460, loss: 0.16250306367874146
step: 470, loss: 0.01325317658483982
step: 480, loss: 0.008089442737400532
step: 490, loss: 0.006637490354478359
step: 500, loss: 0.07689075171947479
step: 510, loss: 0.012480417266488075
step: 520, loss: 0.028876524418592453
step: 530, loss: 0.03062739036977291
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.9181395348837208, f1=0.92243381328379,
best_f1=0.9174311926605505
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.099002406001091
step: 10, loss: 0.006090525072067976
step: 20, loss: 0.012425798922777176
step: 30, loss: 0.014055288396775723

```



step: 40, loss: 0.0026267992798238993  
step: 50, loss: 0.028809132054448128  
step: 60, loss: 0.0024262990336865187  
step: 70, loss: 0.001334786880761385  
step: 80, loss: 0.007312581408768892  
step: 90, loss: 0.017806220799684525  
step: 100, loss: 0.05903849005699158  
step: 110, loss: 0.003600949654355645  
step: 120, loss: 0.022355658933520317  
step: 130, loss: 0.008505878038704395  
step: 140, loss: 0.10756151378154755  
step: 150, loss: 0.003220140468329191  
step: 160, loss: 0.007989615201950073  
step: 170, loss: 0.005529308225959539  
step: 180, loss: 0.015522211790084839  
step: 190, loss: 0.0033073294907808304  
step: 200, loss: 0.0032369475811719894  
step: 210, loss: 0.12735871970653534  
step: 220, loss: 0.010082060471177101  
step: 230, loss: 0.2759178876876831  
step: 240, loss: 0.02955695614218712  
step: 250, loss: 0.02448606677353382  
step: 260, loss: 0.12124388664960861  
step: 270, loss: 0.1410471647977829  
step: 280, loss: 0.0058372593484818935  
step: 290, loss: 0.10795409232378006  
step: 300, loss: 0.02660749852657318  
step: 310, loss: 0.025119638070464134  
step: 320, loss: 0.08124293386936188  
step: 330, loss: 0.19215713441371918  
step: 340, loss: 0.06610709428787231  
step: 350, loss: 0.005430279765278101  
step: 360, loss: 0.006820241920650005  
step: 370, loss: 0.016765573993325233  
step: 380, loss: 0.00242931698448956  
step: 390, loss: 0.08018196374177933  
step: 400, loss: 0.022875986993312836  
step: 410, loss: 0.2763610780239105  
step: 420, loss: 0.011260065250098705  
step: 430, loss: 0.04072489216923714  
step: 440, loss: 0.08348660171031952  
step: 450, loss: 0.016876667737960815  
step: 460, loss: 0.011125387623906136  
step: 470, loss: 0.018638670444488525  
step: 480, loss: 0.004386304412037134  
step: 490, loss: 0.1350550502538681  
step: 500, loss: 0.049441635608673096  
step: 510, loss: 0.05000846087932587

```

step: 520, loss: 0.011502704583108425
step: 530, loss: 0.0009887501364573836
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.9234997709573981, f1=0.921173235563703,
best_f1=0.921173235563703
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007366399746388197
step: 10, loss: 0.008690385147929192
step: 20, loss: 0.017466438934206963
step: 30, loss: 0.0031231355387717485
step: 40, loss: 0.024266939610242844
step: 50, loss: 0.03269375115633011
step: 60, loss: 0.01855562999844551
step: 70, loss: 0.006791498512029648
step: 80, loss: 0.0031931486446410418
step: 90, loss: 0.010920301079750061
step: 100, loss: 0.004449906758964062
step: 110, loss: 0.0018588397651910782
step: 120, loss: 0.016647877171635628
step: 130, loss: 0.004124365746974945
step: 140, loss: 0.0032246499322354794
step: 150, loss: 0.0023192823864519596
step: 160, loss: 0.20363792777061462
step: 170, loss: 0.01943323016166687
step: 180, loss: 0.01125867199152708
step: 190, loss: 0.001171266078017652
step: 200, loss: 0.0011923174606636167
step: 210, loss: 0.0031187175773084164
step: 220, loss: 0.0007770682568661869
step: 230, loss: 0.0008732335991226137
step: 240, loss: 0.02179674431681633
step: 250, loss: 0.0006300812237896025
step: 260, loss: 0.004255691077560186
step: 270, loss: 0.002407823922112584
step: 280, loss: 0.04760251194238663
step: 290, loss: 0.0026979355607181787
step: 300, loss: 0.11972230672836304
step: 310, loss: 0.011095833964645863
step: 320, loss: 0.02220776677131653
step: 330, loss: 0.01629921980202198

```

```

step: 340, loss: 0.14108774065971375
step: 350, loss: 0.05249796435236931
step: 360, loss: 0.04694170877337456
step: 370, loss: 0.0043024942278862
step: 380, loss: 0.07415325194597244
step: 390, loss: 0.009960385039448738
step: 400, loss: 0.011886972934007645
step: 410, loss: 0.0009720420348457992
step: 420, loss: 0.004161288030445576
step: 430, loss: 0.006046981085091829
step: 440, loss: 0.0024751867167651653
step: 450, loss: 0.05816977843642235
step: 460, loss: 0.0025899475440382957
step: 470, loss: 0.038708243519067764
step: 480, loss: 0.00986423622816801
step: 490, loss: 0.002940586069598794
step: 500, loss: 0.12077760696411133
step: 510, loss: 0.0057778325863182545
step: 520, loss: 0.0008670812821947038
step: 530, loss: 0.00983031839132309
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.926012098650535, f1=0.926012098650535,
best_f1=0.926012098650535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015283001121133566
step: 10, loss: 0.09626718610525131
step: 20, loss: 0.003707477357238531
step: 30, loss: 0.009840268641710281
step: 40, loss: 0.0052741048857569695
step: 50, loss: 0.004844087176024914
step: 60, loss: 0.0018251454457640648
step: 70, loss: 0.005627289414405823
step: 80, loss: 0.006127225235104561
step: 90, loss: 0.04352548345923424
step: 100, loss: 0.06705311685800552
step: 110, loss: 0.06309893727302551
step: 120, loss: 0.008790508843958378
step: 130, loss: 0.0009573674178682268
step: 140, loss: 0.021845050156116486
step: 150, loss: 0.0015588179230690002

```

```

step: 160, loss: 0.006699404679238796
step: 170, loss: 0.0001505030522821471
step: 180, loss: 0.003937728237360716
step: 190, loss: 0.01497820857912302
step: 200, loss: 0.0007848451496101916
step: 210, loss: 0.004241915885359049
step: 220, loss: 0.0007286889012902975
step: 230, loss: 0.002938153687864542
step: 240, loss: 0.005682648625224829
step: 250, loss: 0.0017667601350694895
step: 260, loss: 0.005144804250448942
step: 270, loss: 0.0023286130744963884
step: 280, loss: 0.00501586589962244
step: 290, loss: 0.0005811623996123672
step: 300, loss: 0.0016605721320956945
step: 310, loss: 0.006624961271882057
step: 320, loss: 0.021323474124073982
step: 330, loss: 0.0007507162517867982
step: 340, loss: 0.00037196153425611556
step: 350, loss: 0.033036429435014725
step: 360, loss: 0.0007168400334194303
step: 370, loss: 0.0028653377667069435
step: 380, loss: 0.012008348479866982
step: 390, loss: 0.05025200545787811
step: 400, loss: 0.0024415815714746714
step: 410, loss: 0.0002092206123052165
step: 420, loss: 0.013507621362805367
step: 430, loss: 0.00381889077834785
step: 440, loss: 0.020504305139183998
step: 450, loss: 0.008097739890217781
step: 460, loss: 0.11077176779508591
step: 470, loss: 0.02785356156527996
step: 480, loss: 0.009563049301505089
step: 490, loss: 0.0011079220566898584
step: 500, loss: 0.0018276215996593237
step: 510, loss: 0.0004381290345918387
step: 520, loss: 0.016647882759571075
step: 530, loss: 0.0020426386035978794
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9167446211412534, f1=0.9187935034802784,
best_f1=0.926012098650535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04143817722797394
step: 10, loss: 0.057486891746520996

```

step: 20, loss: 0.0030915301758795977  
step: 30, loss: 0.0038142583798617125  
step: 40, loss: 0.03551466390490532  
step: 50, loss: 0.004349351394921541  
step: 60, loss: 0.03437548875808716  
step: 70, loss: 0.009468832984566689  
step: 80, loss: 0.006014480721205473  
step: 90, loss: 0.00012607969983946532  
step: 100, loss: 0.0016550121363252401  
step: 110, loss: 0.00169146538246423  
step: 120, loss: 0.0022309222258627415  
step: 130, loss: 0.0005998499691486359  
step: 140, loss: 0.011692182160913944  
step: 150, loss: 9.231131843989715e-05  
step: 160, loss: 0.0010278847767040133  
step: 170, loss: 0.0039880587719380856  
step: 180, loss: 0.012353766709566116  
step: 190, loss: 0.0006481348536908627  
step: 200, loss: 0.003749803639948368  
step: 210, loss: 0.0016852265689522028  
step: 220, loss: 0.07016659528017044  
step: 230, loss: 0.016548415645956993  
step: 240, loss: 0.008649857714772224  
step: 250, loss: 0.0018785524880513549  
step: 260, loss: 0.008068418130278587  
step: 270, loss: 0.0013521990040317178  
step: 280, loss: 0.03642170876264572  
step: 290, loss: 0.009571452625095844  
step: 300, loss: 0.00025290180929005146  
step: 310, loss: 0.000982905738055706  
step: 320, loss: 0.0385127067565918  
step: 330, loss: 0.001198251498863101  
step: 340, loss: 0.2416364997625351  
step: 350, loss: 0.002842491492629051  
step: 360, loss: 0.036369241774082184  
step: 370, loss: 0.061796776950359344  
step: 380, loss: 0.0075115542858839035  
step: 390, loss: 0.001058029942214489  
step: 400, loss: 0.004730776883661747  
step: 410, loss: 0.0015720488736405969  
step: 420, loss: 0.0014309618854895234  
step: 430, loss: 0.006730163004249334  
step: 440, loss: 0.0033662039786577225  
step: 450, loss: 0.0021726235281676054  
step: 460, loss: 0.00717024365440011  
step: 470, loss: 0.12863081693649292  
step: 480, loss: 0.004365719389170408  
step: 490, loss: 0.0009427547338418663

```

step: 500, loss: 0.00030533241806551814
step: 510, loss: 0.0027536635752767324
step: 520, loss: 0.07097137719392776
step: 530, loss: 0.01883813366293907
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9225092250922509, f1=0.9212962962962964,
best_f1=0.926012098650535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07821223884820938
step: 10, loss: 0.016850102692842484
step: 20, loss: 0.009360521100461483
step: 30, loss: 0.020237458869814873
step: 40, loss: 0.0022346002515405416
step: 50, loss: 0.0025250371545553207
step: 60, loss: 0.0006506580975838006
step: 70, loss: 0.022140726447105408
step: 80, loss: 0.004830429796129465
step: 90, loss: 9.03145846677944e-05
step: 100, loss: 0.0003073278348892927
step: 110, loss: 0.0008687057415954769
step: 120, loss: 0.008617108687758446
step: 130, loss: 0.00010852843115571886
step: 140, loss: 4.286945113562979e-05
step: 150, loss: 0.0010644085705280304
step: 160, loss: 0.0043214550241827965
step: 170, loss: 0.003950290381908417
step: 180, loss: 0.0004610467585735023
step: 190, loss: 0.0008626288617961109
step: 200, loss: 0.007471973076462746
step: 210, loss: 0.0010807141661643982
step: 220, loss: 0.018410993739962578
step: 230, loss: 0.009737621992826462
step: 240, loss: 0.021907256916165352
step: 250, loss: 0.006628527771681547
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 260, loss: 0.11195141077041626
step: 270, loss: 0.0002541945141274482
step: 280, loss: 0.0015019611455500126
step: 290, loss: 0.07938173413276672
step: 300, loss: 0.0009030196815729141
step: 310, loss: 0.04690765589475632
step: 320, loss: 0.00042105591273866594
step: 330, loss: 0.053586505353450775
step: 340, loss: 0.000409014115575701

```

```
step: 350, loss: 0.002938993275165558
step: 360, loss: 0.016807949170470238
step: 370, loss: 0.0010258691618219018
step: 380, loss: 0.0030432736966758966
step: 390, loss: 0.00017147658218163997
step: 400, loss: 0.05904421955347061
step: 410, loss: 0.0013914275914430618
step: 420, loss: 0.0021186743397265673
step: 430, loss: 0.0020950022153556347
step: 440, loss: 0.000397656811401248
step: 450, loss: 0.00037591144791804254
step: 460, loss: 0.0008251451072283089
step: 470, loss: 0.0001035826062434353
step: 480, loss: 0.003663018811494112
step: 490, loss: 0.0011733619030565023
step: 500, loss: 0.0005629113875329494
step: 510, loss: 0.00045813259202986956
step: 520, loss: 0.0006615514867007732
step: 530, loss: 0.00040873276884667575
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9210406207211318, f1=0.9243316719528771,
best_f1=0.926012098650535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00035047103301621974
step: 10, loss: 0.013779289089143276
step: 20, loss: 0.0018354430794715881
step: 30, loss: 0.0003727362200152129
step: 40, loss: 0.13941460847854614
step: 50, loss: 0.0045922547578811646
step: 60, loss: 0.0005050786421634257
step: 70, loss: 0.10772980004549026
step: 80, loss: 0.0003451041702646762
step: 90, loss: 0.0013308385387063026
step: 100, loss: 0.00016538605268578976
step: 110, loss: 0.0016377524007111788
step: 120, loss: 0.06991495192050934
step: 130, loss: 0.006180538330227137
step: 140, loss: 0.03378372639417648
step: 150, loss: 0.0001776607969077304
step: 160, loss: 0.0022223624400794506
step: 170, loss: 0.0004907854599878192
step: 180, loss: 0.000275046011665836
step: 190, loss: 0.009300614707171917
step: 200, loss: 0.0003313080233056098
```

```

step: 210, loss: 0.0012147321831434965
step: 220, loss: 0.0029999185353517532
step: 230, loss: 0.00016878168389666826
step: 240, loss: 0.002115347422659397
step: 250, loss: 0.0009075740235857666
step: 260, loss: 0.00012308463919907808
step: 270, loss: 0.01710522174835205
step: 280, loss: 0.00014388909039553255
step: 290, loss: 8.195418195100501e-05
step: 300, loss: 0.12390384078025818
step: 310, loss: 0.0003172785509377718
step: 320, loss: 0.0015103643527254462
step: 330, loss: 0.00019729083578567952
step: 340, loss: 0.0006997823948040605
step: 350, loss: 0.06472596526145935
step: 360, loss: 0.025193817913532257
step: 370, loss: 0.00024194034631364048
step: 380, loss: 8.263226482085884e-05
step: 390, loss: 0.0003466581110842526
step: 400, loss: 0.004015388432890177
step: 410, loss: 7.14758934918791e-05
step: 420, loss: 0.0014770464040338993
step: 430, loss: 5.869617234566249e-05
step: 440, loss: 0.0013266261667013168
step: 450, loss: 0.00038292352110147476
step: 460, loss: 0.0023793024010956287
step: 470, loss: 6.95832641213201e-05
step: 480, loss: 0.05087381601333618
step: 490, loss: 0.008798224851489067
step: 500, loss: 0.0017396926414221525
step: 510, loss: 0.010695649310946465
step: 520, loss: 0.00011548698967089877
step: 530, loss: 0.0003158013860229403
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9180176007410837, f1=0.92619926199262,
best_f1=0.926012098650535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009311094181612134
step: 10, loss: 5.800217331852764e-05
step: 20, loss: 0.00011583173181861639
step: 30, loss: 0.00011752817954402417
step: 40, loss: 0.003819308942183852
step: 50, loss: 0.0002137806877726689
step: 60, loss: 0.00036522187292575836

```



step: 70, loss: 0.0011061878176406026  
step: 80, loss: 0.00912882573902607  
step: 90, loss: 0.0021472200751304626  
step: 100, loss: 0.00014027216820977628  
step: 110, loss: 0.00107545533683151  
step: 120, loss: 0.09025867283344269  
step: 130, loss: 0.0002610345254652202  
step: 140, loss: 0.0008864307892508805  
step: 150, loss: 0.023369470611214638  
step: 160, loss: 0.0005931798950769007  
step: 170, loss: 0.00010671903874026611  
step: 180, loss: 0.017499571666121483  
step: 190, loss: 0.0014502268750220537  
step: 200, loss: 0.00039683215436525643  
step: 210, loss: 5.995236278977245e-05  
step: 220, loss: 0.00015580785111524165  
step: 230, loss: 0.0006063871551305056  
step: 240, loss: 9.594712173566222e-05  
step: 250, loss: 0.0002937503741122782  
step: 260, loss: 0.002268451265990734  
step: 270, loss: 0.0015746040735393763  
step: 280, loss: 3.5940764064434916e-05  
step: 290, loss: 0.0005932586500421166  
step: 300, loss: 0.004274439532309771  
step: 310, loss: 0.00025188224390149117  
step: 320, loss: 0.0004626297450158745  
step: 330, loss: 0.00017152438522316515  
step: 340, loss: 0.0025281780399382114  
step: 350, loss: 0.00018339008965995163  
step: 360, loss: 0.011194090358912945  
step: 370, loss: 0.0007805681671015918  
step: 380, loss: 0.00506251584738493  
step: 390, loss: 0.0011779994238168001  
step: 400, loss: 0.007850014604628086  
step: 410, loss: 0.0010950721334666014  
step: 420, loss: 0.010150684975087643  
step: 430, loss: 0.0020605733152478933  
step: 440, loss: 0.00016743595188017935  
step: 450, loss: 0.003979848697781563  
step: 460, loss: 0.004343593493103981  
step: 470, loss: 0.0009096356225199997  
step: 480, loss: 0.001142452354542911  
step: 490, loss: 0.046185046434402466  
step: 500, loss: 0.007702725939452648  
step: 510, loss: 8.30353528726846e-05  
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0  
step: 520, loss: 0.002422158606350422  
step: 530, loss: 0.00058560207253322

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9235697940503432, f1=0.9275626423690205,
best_f1=0.926012098650535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005021713441237807
step: 10, loss: 0.0020421280059963465
step: 20, loss: 3.8085971027612686e-05
step: 30, loss: 0.0010476900497451425
step: 40, loss: 0.001300301868468523
step: 50, loss: 0.0011356357717886567
step: 60, loss: 0.001707383431494236
step: 70, loss: 0.00010585134441498667
step: 80, loss: 0.0008243293850682676
step: 90, loss: 0.007101903669536114
step: 100, loss: 0.00020034490444231778
step: 110, loss: 0.001670737867243588
step: 120, loss: 0.006225548684597015
step: 130, loss: 0.0004096741322427988
step: 140, loss: 0.00024154773564077914
step: 150, loss: 0.00021693216694984585
step: 160, loss: 0.00030592584516853094
step: 170, loss: 0.08366284519433975
step: 180, loss: 6.951463728910312e-05
step: 190, loss: 0.028987867757678032
step: 200, loss: 0.01889691688120365
step: 210, loss: 0.00036818961962126195
step: 220, loss: 0.00024292911984957755
step: 230, loss: 0.21283259987831116
step: 240, loss: 0.0014635395491495728
step: 250, loss: 0.00015647411055397242
step: 260, loss: 3.3351414458593354e-05
step: 270, loss: 0.0016398170264437795
step: 280, loss: 0.0011916401563212276
step: 290, loss: 0.014865882694721222
step: 300, loss: 0.0038353861309587955
step: 310, loss: 0.0017706131329759955
step: 320, loss: 0.0145891597494483
step: 330, loss: 0.00037294349749572575
step: 340, loss: 0.00019963132217526436
step: 350, loss: 0.002480280352756381
step: 360, loss: 0.00011346893734298646
step: 370, loss: 3.4970576962223276e-05
step: 380, loss: 7.735128019703552e-05
step: 390, loss: 0.0017047086730599403
```

```

step: 400, loss: 6.640122592216358e-05
step: 410, loss: 0.0036989350337535143
step: 420, loss: 0.001742960070259869
step: 430, loss: 0.015627827495336533
step: 440, loss: 0.0001538138894829899
step: 450, loss: 0.0006467251805588603
step: 460, loss: 0.0017045637359842658
step: 470, loss: 0.0005518798716366291
step: 480, loss: 0.0006491130334325135
step: 490, loss: 0.001414807396940887
step: 500, loss: 0.0005615822738036513
step: 510, loss: 0.0367855541408062
step: 520, loss: 3.875504989991896e-05
step: 530, loss: 0.00013482193753588945
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9167455061494796, f1=0.9288742345737163,
best_f1=0.926012098650535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00048112260992638767
step: 10, loss: 0.00012753297050949186
step: 20, loss: 2.1621279302053154e-05
step: 30, loss: 3.564991129678674e-05
step: 40, loss: 8.368244743905962e-05
step: 50, loss: 0.019826648756861687
step: 60, loss: 0.0011353036388754845
step: 70, loss: 0.00015690758300479501
step: 80, loss: 0.0016230256296694279
step: 90, loss: 0.0010909572010859847
step: 100, loss: 0.00021962735627312213
step: 110, loss: 0.002635408192873001
step: 120, loss: 0.0007506885449402034
step: 130, loss: 0.014730837196111679
step: 140, loss: 4.494632958085276e-05
step: 150, loss: 0.0003223420644644648
step: 160, loss: 0.0138253103941679
step: 170, loss: 0.0056376042775809765
step: 180, loss: 0.0003381433489266783
step: 190, loss: 5.571973088080995e-05
step: 200, loss: 0.00025658932281658053
step: 210, loss: 0.0006321955006569624
step: 220, loss: 0.0002916727389674634
step: 230, loss: 0.003561911638826132
step: 240, loss: 0.00038613940705545247
step: 250, loss: 0.00022818674915470183

```

```

step: 260, loss: 0.0008594990940764546
step: 270, loss: 0.0064455559477210045
step: 280, loss: 0.00048072912613861263
step: 290, loss: 0.013542178086936474
step: 300, loss: 0.005478335078805685
step: 310, loss: 0.007358220871537924
step: 320, loss: 0.03248066455125809
step: 330, loss: 0.0001400880137225613
step: 340, loss: 0.0024026299361139536
step: 350, loss: 0.015596914105117321
step: 360, loss: 0.0008761680219322443
step: 370, loss: 8.928278839448467e-05
step: 380, loss: 0.00040875421836972237
step: 390, loss: 6.170448614284396e-05
step: 400, loss: 0.0004345702182035893
step: 410, loss: 0.010074220597743988
step: 420, loss: 0.007359550334513187
step: 430, loss: 0.0007693818188272417
step: 440, loss: 0.001649752608500421
step: 450, loss: 0.01744028925895691
step: 460, loss: 0.0023692601826041937
step: 470, loss: 0.0009772905614227057
step: 480, loss: 0.0009079510346055031
step: 490, loss: 0.0006508249789476395
step: 500, loss: 0.00311212963424623
step: 510, loss: 0.00023007747950032353
step: 520, loss: 0.021337315440177917
step: 530, loss: 0.00015081030142027885
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.923216719672876, f1=0.9255079006772009,
best_f1=0.926012098650535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003332897322252393
step: 10, loss: 0.0007171517936512828
step: 20, loss: 0.02783765271306038
step: 30, loss: 0.0005289409891702235
step: 40, loss: 0.0001888496772153303
step: 50, loss: 0.00013582213432528079
step: 60, loss: 0.0002558245323598385
step: 70, loss: 8.455581701127812e-05
step: 80, loss: 2.2954824089538306e-05
step: 90, loss: 0.00038446622784249485
step: 100, loss: 0.00012784081627614796
step: 110, loss: 0.000817509600892663

```

```
step: 120, loss: 0.0007127815624698997
step: 130, loss: 0.0008972485084086657
step: 140, loss: 0.001081668771803379
step: 150, loss: 0.00044458432239480317
step: 160, loss: 0.00028011095128022134
step: 170, loss: 0.010950748808681965
step: 180, loss: 6.148456304799765e-05
step: 190, loss: 0.003054311266168952
step: 200, loss: 0.009129329584538937
step: 210, loss: 0.00314153078943491
step: 220, loss: 0.00012978946324437857
step: 230, loss: 0.0001928699348354712
step: 240, loss: 0.003197225509211421
step: 250, loss: 0.05728374049067497
step: 260, loss: 0.00023934045748319477
step: 270, loss: 0.014276527799665928
step: 280, loss: 4.765360790770501e-05
step: 290, loss: 0.006129689514636993
step: 300, loss: 0.0006135731237009168
step: 310, loss: 0.0005781284999102354
step: 320, loss: 2.7670346753438935e-05
step: 330, loss: 0.002153342356905341
step: 340, loss: 0.0005519986152648926
step: 350, loss: 0.022500187158584595
step: 360, loss: 7.07623694324866e-05
step: 370, loss: 7.353659748332575e-05
step: 380, loss: 0.0021097625140100718
step: 390, loss: 0.0005980432615615427
step: 400, loss: 0.0007174999336712062
step: 410, loss: 0.0003296168288215995
step: 420, loss: 0.0005380703951232135
step: 430, loss: 0.00034209733712486923
step: 440, loss: 0.0006401445134542882
step: 450, loss: 0.0007856212905608118
step: 460, loss: 2.213118932559155e-05
step: 470, loss: 0.022363577038049698
step: 480, loss: 0.0009573693969286978
step: 490, loss: 0.0011881985701620579
step: 500, loss: 0.002970263594761491
step: 510, loss: 6.0657403082586825e-05
step: 520, loss: 0.00013214022328611463
step: 530, loss: 9.312056499766186e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.923646459972235, f1=0.929368029739777,
best_f1=0.926012098650535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.004965623375028372
step: 10, loss: 0.001742711290717125
step: 20, loss: 3.998796455562115e-05
step: 30, loss: 0.00022662643459625542
step: 40, loss: 0.0008910408359952271
step: 50, loss: 6.0434704209910706e-05
step: 60, loss: 2.6426221666042693e-05
step: 70, loss: 8.08496042736806e-05
step: 80, loss: 0.003828148590400815
step: 90, loss: 0.001270591514185071
step: 100, loss: 0.0001976941421162337
step: 110, loss: 0.00012229851563461125
step: 120, loss: 0.0019044260261580348
step: 130, loss: 0.00019550863362383097
step: 140, loss: 0.01708214357495308
step: 150, loss: 0.0008272668928839266
step: 160, loss: 0.00037942748167552054
step: 170, loss: 0.02074461616575718
step: 180, loss: 0.00542102986946702
step: 190, loss: 0.000191794810234569
step: 200, loss: 0.0010772718815132976
step: 210, loss: 0.0007075038156472147
step: 220, loss: 0.0004753278335556388
step: 230, loss: 2.829268305504229e-05
step: 240, loss: 0.001604854129254818
step: 250, loss: 2.041421612375416e-05
step: 260, loss: 9.722991671878844e-05
step: 270, loss: 0.002444472862407565
step: 280, loss: 0.00017987246974371374
step: 290, loss: 7.276092219399288e-05
step: 300, loss: 5.78367144044023e-05
step: 310, loss: 0.00013236593804322183
step: 320, loss: 0.00015116191934794188
step: 330, loss: 0.0003157332830596715
step: 340, loss: 4.4972370233153924e-05
step: 350, loss: 0.001025466714054346
step: 360, loss: 0.014828067272901535
step: 370, loss: 6.731179018970579e-05
step: 380, loss: 0.00025346726761199534
step: 390, loss: 0.00169442780315876
step: 400, loss: 4.124502811464481e-05
step: 410, loss: 0.00042729126289486885
step: 420, loss: 0.00019909524417016655
step: 430, loss: 0.0002604894689284265
step: 440, loss: 2.2935812012292445e-05
step: 450, loss: 0.0001306196063524112
```

```

step: 460, loss: 0.00086695805657655
step: 470, loss: 0.003326678415760398
step: 480, loss: 3.687504431582056e-05
step: 490, loss: 0.00011647643987089396
step: 500, loss: 0.0008140582358464599
step: 510, loss: 0.002149266889318824
step: 520, loss: 0.0009691667510196567
step: 530, loss: 2.6068557417602278e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.922792417938049, f1=0.9291338582677167,
best_f1=0.926012098650535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.6258511904161423e-05
step: 10, loss: 0.0002566805051174015
step: 20, loss: 0.0004643838037736714
step: 30, loss: 0.0007715283427387476
step: 40, loss: 5.931374471401796e-05
step: 50, loss: 8.755783346714452e-05
step: 60, loss: 3.1909061362966895e-05
step: 70, loss: 0.00010916755854850635
step: 80, loss: 0.0004901587381027639
step: 90, loss: 1.5724293916719034e-05
step: 100, loss: 3.288501829956658e-05
step: 110, loss: 0.0200165044516325
step: 120, loss: 0.00015349291788879782
step: 130, loss: 0.00018337917572353035
step: 140, loss: 2.719608528423123e-05
step: 150, loss: 0.010470657609403133
step: 160, loss: 0.0001928943529492244
step: 170, loss: 0.00025797318085096776
step: 180, loss: 0.0004398193850647658
step: 190, loss: 0.0007017969037406147
step: 200, loss: 4.1487248381599784e-05
step: 210, loss: 0.001488313777372241
step: 220, loss: 0.00020545252482406795
step: 230, loss: 0.004263265058398247
step: 240, loss: 4.886267561232671e-05
step: 250, loss: 8.822459494695067e-05
step: 260, loss: 0.001112492522224784
step: 270, loss: 0.007620854303240776
step: 280, loss: 0.00028859719168394804
step: 290, loss: 0.000411500979680568
step: 300, loss: 0.0003776295343413949
step: 310, loss: 0.0003401116409804672

```

```

step: 320, loss: 4.736950359074399e-05
step: 330, loss: 0.0007187516894191504
step: 340, loss: 0.005488025490194559
step: 350, loss: 0.0006977865705266595
step: 360, loss: 0.0006208349950611591
step: 370, loss: 2.0626473997253925e-05
step: 380, loss: 6.769697210984305e-05
step: 390, loss: 0.007677861489355564
step: 400, loss: 1.930391590576619e-05
step: 410, loss: 4.979280129191466e-05
step: 420, loss: 7.990564336068928e-05
step: 430, loss: 4.379094752948731e-05
step: 440, loss: 0.000314958393573761
step: 450, loss: 0.0006405989988707006
step: 460, loss: 0.00010336288687540218
step: 470, loss: 0.00015948554209899157
step: 480, loss: 7.267174805747345e-05
step: 490, loss: 0.00021233016741462052
step: 500, loss: 0.0006827234174124897
step: 510, loss: 0.16881033778190613
step: 520, loss: 7.896679016994312e-05
step: 530, loss: 1.665931449679192e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9237918215613383, f1=0.9307942405945193,
best_f1=0.926012098650535
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic

```



```

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:14, 395.57it/s]
load_f1 = 0.9246919214970334
real_f1 = 0.9234972677595629
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:10, 402.93it/s]

```

#### 4.1.4 Amazon-Google - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
    --task Structured/Amazon-Google \
    --batch_size 32 \
    --max_len 256 \
    --lr 3e-5 \
    --n_epochs 15 \
    --finetuning \
    --lm distilbert \
    --fp16 \
    --save_model \
    --dk general \
    --dk product \
    --summarize \
    --da entry_swap \
    --da attr_del

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Amazon-Google.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
    --task Structured/Amazon-Google \

```

```

--input_path input/input_small.jsonl \
--output_path output/output_small.jsonl \
--lm distilbert \
--use_gpu \
--fp16 \
--checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

warnings.warn("An input tensor was not cuda.")

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0 /usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`. In PyTorch 1.1.0 and later, you should call them in the opposite order:

`optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

"https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate", UserWarning)

step: 0, loss: 0.8279861211776733

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

step: 10, loss: 0.0601588636636734

step: 20, loss: 0.38364657759666443

```

step: 30, loss: 0.37353771924972534
step: 40, loss: 0.5152801871299744
step: 50, loss: 0.30256322026252747
step: 60, loss: 0.3670318126678467
step: 70, loss: 0.2654430866241455
step: 80, loss: 0.2668716609477997
step: 90, loss: 0.3964742422103882
step: 100, loss: 0.21630078554153442
step: 110, loss: 0.33346283435821533
step: 120, loss: 0.23120790719985962
step: 130, loss: 0.2334796041250229
step: 140, loss: 0.2632446587085724
step: 150, loss: 0.29181671142578125
step: 160, loss: 0.18334800004959106
step: 170, loss: 0.1713995337486267
step: 180, loss: 0.2339101880788803
step: 190, loss: 0.1665123552083969
step: 200, loss: 0.17243646085262299
step: 210, loss: 0.48746809363365173
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.46625766871165647, f1=0.5415778251599146,
best_f1=0.5415778251599146
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09468042105436325
step: 10, loss: 0.10343126952648163
step: 20, loss: 0.3148369789123535
step: 30, loss: 0.16162073612213135
step: 40, loss: 0.11067128926515579
step: 50, loss: 0.2104109823703766
step: 60, loss: 0.0707300528883934
step: 70, loss: 0.142543762922287
step: 80, loss: 0.21382033824920654
step: 90, loss: 0.25242263078689575
step: 100, loss: 0.11251436918973923
step: 110, loss: 0.05025683343410492
step: 120, loss: 0.3221568167209625
step: 130, loss: 0.22862371802330017
step: 140, loss: 0.262624055147171
step: 150, loss: 0.22063307464122772
step: 160, loss: 0.1385446935892105

```

```

step: 170, loss: 0.17898397147655487
step: 180, loss: 0.2622257173061371
step: 190, loss: 0.1889907419681549
step: 200, loss: 0.16352204978466034
step: 210, loss: 0.2581196427345276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.5779334500875657, f1=0.6135181975736568,
best_f1=0.6135181975736568
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2558578550815582
step: 10, loss: 0.22422869503498077
step: 20, loss: 0.27978044748306274
step: 30, loss: 0.04240266978740692
step: 40, loss: 0.25380030274391174
step: 50, loss: 0.08569539338350296
step: 60, loss: 0.4092632532119751
step: 70, loss: 0.09245312958955765
step: 80, loss: 0.08840673416852951
step: 90, loss: 0.17501436173915863
step: 100, loss: 0.031603772193193436
step: 110, loss: 0.17174910008907318
step: 120, loss: 0.25433528423309326
step: 130, loss: 0.10053660720586777
step: 140, loss: 0.17877577245235443
step: 150, loss: 0.3837451636791229
step: 160, loss: 0.0780266746878624
step: 170, loss: 0.16300025582313538
step: 180, loss: 0.10178528726100922
step: 190, loss: 0.2246822863817215
step: 200, loss: 0.18929465115070343
step: 210, loss: 0.2219226360321045
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.5897435897435896, f1=0.6394052044609665,
best_f1=0.6394052044609665

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09316341578960419
step: 10, loss: 0.056940000504255295
step: 20, loss: 0.09877682477235794
step: 30, loss: 0.032438118010759354
step: 40, loss: 0.023695949465036392
step: 50, loss: 0.0911785364151001
step: 60, loss: 0.03588211536407471
step: 70, loss: 0.2459549903869629
step: 80, loss: 0.14929328858852386
step: 90, loss: 0.010499910451471806
step: 100, loss: 0.08847910165786743
step: 110, loss: 0.23738576471805573
step: 120, loss: 0.10001440346240997
step: 130, loss: 0.2519236207008362
step: 140, loss: 0.14367753267288208
step: 150, loss: 0.09304483234882355
step: 160, loss: 0.0891294926404953
step: 170, loss: 0.16379143297672272
step: 180, loss: 0.15227994322776794
step: 190, loss: 0.08542278409004211
step: 200, loss: 0.10421028733253479
step: 210, loss: 0.06269983947277069
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.5637583892617449, f1=0.5633333333333332,
best_f1=0.6394052044609665
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09796475619077682
step: 10, loss: 0.09088146686553955
step: 20, loss: 0.048901885747909546
step: 30, loss: 0.07813076674938202
step: 40, loss: 0.30183112621307373
step: 50, loss: 0.12124928086996078
step: 60, loss: 0.028795428574085236
step: 70, loss: 0.1050066351890564
step: 80, loss: 0.07371628284454346
step: 90, loss: 0.0733482763171196
step: 100, loss: 0.08449137210845947
step: 110, loss: 0.007438701577484608
step: 120, loss: 0.0439843125641346
step: 130, loss: 0.07266897708177567
step: 140, loss: 0.034134525805711746

```

```

step: 150, loss: 0.201504647731781
step: 160, loss: 0.04130222648382187
step: 170, loss: 0.13384588062763214
step: 180, loss: 0.17314355075359344
step: 190, loss: 0.13798931241035461
step: 200, loss: 0.11723994463682175
step: 210, loss: 0.04330454021692276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.5826771653543307, f1=0.5979797979797981,
best_f1=0.6394052044609665
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06320303678512573
step: 10, loss: 0.02566310204565525
step: 20, loss: 0.08578938990831375
step: 30, loss: 0.2547593116760254
step: 40, loss: 0.02896377071738243
step: 50, loss: 0.08991727232933044
step: 60, loss: 0.11921725422143936
step: 70, loss: 0.0077166748233139515
step: 80, loss: 0.18442587554454803
step: 90, loss: 0.07749643176794052
step: 100, loss: 0.02111946791410446
step: 110, loss: 0.03774989768862724
step: 120, loss: 0.0305409487336874
step: 130, loss: 0.11017744243144989
step: 140, loss: 0.14126808941364288
step: 150, loss: 0.014829843305051327
step: 160, loss: 0.07206039130687714
step: 170, loss: 0.01229240745306015
step: 180, loss: 0.0034438876900821924
step: 190, loss: 0.06340785324573517
step: 200, loss: 0.023264123126864433
step: 210, loss: 0.02935827523469925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.5803108808290156, f1=0.5793103448275861,
best_f1=0.6394052044609665
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07512036710977554
step: 10, loss: 0.03207077085971832
step: 20, loss: 0.01761496812105179

```

```

step: 30, loss: 0.025828616693615913
step: 40, loss: 0.05162561684846878
step: 50, loss: 0.08726506680250168
step: 60, loss: 0.14171704649925232
step: 70, loss: 0.01090623065829277
step: 80, loss: 0.014758139848709106
step: 90, loss: 0.014800065197050571
step: 100, loss: 0.036989931017160416
step: 110, loss: 0.11357370018959045
step: 120, loss: 0.3190372586250305
step: 130, loss: 0.0790700912475586
step: 140, loss: 0.009169647470116615
step: 150, loss: 0.08510176092386246
step: 160, loss: 0.16376133263111115
step: 170, loss: 0.016307612881064415
step: 180, loss: 0.0033300830982625484
step: 190, loss: 0.11644463986158371
step: 200, loss: 0.05045032128691673
step: 210, loss: 0.1675424724817276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.5888689407540395, f1=0.6029411764705882,
best_f1=0.6394052044609665
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04471749812364578
step: 10, loss: 0.01908668503165245
step: 20, loss: 0.08873189985752106
step: 30, loss: 0.05134914070367813
step: 40, loss: 0.038352034986019135
step: 50, loss: 0.0027016091626137495
step: 60, loss: 0.2145560085773468
step: 70, loss: 0.006883283611387014
step: 80, loss: 0.11314565688371658
step: 90, loss: 0.02753659337759018
step: 100, loss: 0.02271042950451374
step: 110, loss: 0.13004577159881592
step: 120, loss: 0.07833388447761536
step: 130, loss: 0.05355983227491379
step: 140, loss: 0.0035506575368344784
step: 150, loss: 0.02717505395412445
step: 160, loss: 0.029616158455610275
step: 170, loss: 0.02603776752948761
step: 180, loss: 0.020335568115115166
step: 190, loss: 0.009788639843463898
step: 200, loss: 0.047541067004203796

```

```

step: 210, loss: 0.20053213834762573
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.5437788018433181, f1=0.5539906103286385,
best_f1=0.6394052044609665
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028119025751948357
step: 10, loss: 0.15422819554805756
step: 20, loss: 0.03213803842663765
step: 30, loss: 0.19053122401237488
step: 40, loss: 0.03869396075606346
step: 50, loss: 0.03686656057834625
step: 60, loss: 0.0536426417529583
step: 70, loss: 0.0087578184902668
step: 80, loss: 0.012642400339245796
step: 90, loss: 0.02292860671877861
step: 100, loss: 0.06360287964344025
step: 110, loss: 0.032206904143095016
step: 120, loss: 0.19003191590309143
step: 130, loss: 0.00656364718452096
step: 140, loss: 0.05717403069138527
step: 150, loss: 0.020564429461956024
step: 160, loss: 0.0015139105962589383
step: 170, loss: 0.016512664034962654
step: 180, loss: 0.01792888529598713
step: 190, loss: 0.12884359061717987
step: 200, loss: 0.02326420322060585
step: 210, loss: 0.1018117144703865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.5601659751037344, f1=0.612691466083151,
best_f1=0.6394052044609665
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016889380291104317
step: 10, loss: 0.002598430262878537
step: 20, loss: 0.08613891899585724
step: 30, loss: 0.1673871874809265
step: 40, loss: 0.15085457265377045
step: 50, loss: 0.0012754785129800439
step: 60, loss: 0.007477065082639456
step: 70, loss: 0.007371135987341404
step: 80, loss: 0.0008844491676427424

```



```

step: 90, loss: 0.19480812549591064
step: 100, loss: 0.0006009895587339997
step: 110, loss: 0.02104368433356285
step: 120, loss: 0.006616478785872459
step: 130, loss: 0.006726439110934734
step: 140, loss: 0.001628025434911251
step: 150, loss: 0.055375467985868454
step: 160, loss: 0.01897282898426056
step: 170, loss: 0.09392958879470825
step: 180, loss: 0.0021662430372089148
step: 190, loss: 0.008541014045476913
step: 200, loss: 0.012770247645676136
step: 210, loss: 0.015261261723935604
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.5569620253164557, f1=0.5871964679911699,
best_f1=0.6394052044609665
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018386781215667725
step: 10, loss: 0.025617986917495728
step: 20, loss: 0.006171587388962507
step: 30, loss: 0.013314414769411087
step: 40, loss: 0.03589671850204468
step: 50, loss: 0.015689771622419357
step: 60, loss: 0.011247367598116398
step: 70, loss: 0.008450482971966267
step: 80, loss: 0.008649014867842197
step: 90, loss: 0.04946881905198097
step: 100, loss: 0.003919222392141819
step: 110, loss: 0.0037911406252533197
step: 120, loss: 0.02062704972922802
step: 130, loss: 0.014767652377486229
step: 140, loss: 0.07037448137998581
step: 150, loss: 0.036156609654426575
step: 160, loss: 0.039832159876823425
step: 170, loss: 0.09401597827672958
step: 180, loss: 0.057957012206315994
step: 190, loss: 0.002748252125456929
step: 200, loss: 0.0010668131290003657
step: 210, loss: 0.001345123746432364
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.5638095238095238, f1=0.588469184890656,
best_f1=0.6394052044609665

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000454755499958992
step: 10, loss: 0.07306907325983047
step: 20, loss: 0.004156971350312233
step: 30, loss: 0.02443874627351761
step: 40, loss: 0.011151888407766819
step: 50, loss: 0.02640310302376747
step: 60, loss: 0.03643462434411049
step: 70, loss: 0.003251191694289446
step: 80, loss: 0.0023802935611456633
step: 90, loss: 0.013175277039408684
step: 100, loss: 0.005035310052335262
step: 110, loss: 0.0016071232967078686
step: 120, loss: 0.002369017107412219
step: 130, loss: 0.03302135318517685
step: 140, loss: 0.057547569274902344
step: 150, loss: 0.0037970503326505423
step: 160, loss: 0.019685691222548485
step: 170, loss: 0.001525376457720995
step: 180, loss: 0.018832819536328316
step: 190, loss: 0.31723248958587646
step: 200, loss: 0.1769827902317047
step: 210, loss: 0.057794515043497086
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.570230607966457, f1=0.5814977973568282,
best_f1=0.6394052044609665
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019195611821487546
step: 10, loss: 0.013872084207832813
step: 20, loss: 0.009765281341969967
step: 30, loss: 0.006228029262274504
step: 40, loss: 0.008923904038965702
step: 50, loss: 0.0024145508650690317
step: 60, loss: 0.0015844021691009402
step: 70, loss: 0.017914339900016785
step: 80, loss: 0.03440646454691887
step: 90, loss: 0.012762876227498055
step: 100, loss: 0.05657839775085449
step: 110, loss: 0.002348017878830433
step: 120, loss: 0.0018830220215022564
step: 130, loss: 0.0021527810022234917
step: 140, loss: 0.012861697934567928

```

```

step: 150, loss: 0.002555776620283723
step: 160, loss: 0.0033297252375632524
step: 170, loss: 0.005415409337729216
step: 180, loss: 0.04491550102829933
step: 190, loss: 0.0011705022770911455
step: 200, loss: 0.0181976817548275
step: 210, loss: 0.007599607575684786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.5637860082304527, f1=0.6004228329809724,
best_f1=0.6394052044609665
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0020988653413951397
step: 10, loss: 0.06550943851470947
step: 20, loss: 0.007836966775357723
step: 30, loss: 0.003161381697282195
step: 40, loss: 0.07012800872325897
step: 50, loss: 0.015346820466220379
step: 60, loss: 0.02894998900592327
step: 70, loss: 0.08004138618707657
step: 80, loss: 0.058890148997306824
step: 90, loss: 0.12123118340969086
step: 100, loss: 0.005053764674812555
step: 110, loss: 0.0061889952048659325
step: 120, loss: 0.08821916580200195
step: 130, loss: 0.04500177875161171
step: 140, loss: 0.0006043490720912814
step: 150, loss: 0.020913679152727127
step: 160, loss: 0.008316345512866974
step: 170, loss: 0.0951911136507988
step: 180, loss: 0.004871068522334099
step: 190, loss: 0.003102569840848446
step: 200, loss: 0.0016746923793107271
step: 210, loss: 0.061195243149995804
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.570873786407767, f1=0.6040816326530613,
best_f1=0.6394052044609665
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008959797210991383
step: 10, loss: 0.006977576296776533
step: 20, loss: 0.0049065793864429

```

```

step: 30, loss: 0.0008045700960792601
step: 40, loss: 0.0012888298369944096
step: 50, loss: 0.0031409882940351963
step: 60, loss: 0.007114647887647152
step: 70, loss: 0.004610641859471798
step: 80, loss: 0.0042740581557154655
step: 90, loss: 0.007245604880154133
step: 100, loss: 0.062395501881837845
step: 110, loss: 0.00092186318943277
step: 120, loss: 0.13398835062980652
step: 130, loss: 0.013599281199276447
step: 140, loss: 0.006213838700205088
step: 150, loss: 0.009030073881149292
step: 160, loss: 0.013553074561059475
step: 170, loss: 0.05415451526641846
step: 180, loss: 0.02089773863554001
step: 190, loss: 0.05724824592471123
step: 200, loss: 0.018314428627490997
step: 210, loss: 0.12235966324806213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.576923076923077, f1=0.6052104208416834,
best_f1=0.6394052044609665
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type       : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type       : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2293it [00:03, 741.55it/s]
load_f1 = 0.5916030534351145
real_f1 = 0.576923076923077
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:10, 407.95it/s]

```

#### 4.1.5 Walmart-Amazon - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/Walmart-Amazon \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm distilbert \
  --fp16 \
  --save_model \
  --dk general \
  --dk product \
  --summarize \
  --da entry_swap \
  --da attr_del

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Walmart-Amazon.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/Walmart-Amazon \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm distilbert \

```

```
--use_gpu \
--fp16 \
--checkpoint_path checkpoints/
```

```
Downloading: 100% 442/442 [00:00<00:00, 561kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 3.84MB/s]
Downloading: 100% 268M/268M [00:04<00:00, 59.1MB/s]
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.
```

Defaults for this optimization level are:

```
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

```
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
```

```
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
```

```
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
```

```
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate
```

```
"https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
```

```
step: 0, loss: 0.8485904932022095
```

```
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
```

```
step: 10, loss: 0.16293080151081085
```

```
step: 20, loss: 0.15474125742912292
```

```

step: 30, loss: 0.5108357071876526
step: 40, loss: 0.27128687500953674
step: 50, loss: 0.3107626140117645
step: 60, loss: 0.3654560148715973
step: 70, loss: 0.18675518035888672
step: 80, loss: 0.5503184199333191
step: 90, loss: 0.24185368418693542
step: 100, loss: 0.22445142269134521
step: 110, loss: 0.23848727345466614
step: 120, loss: 0.41700199246406555
step: 130, loss: 0.3439408838748932
step: 140, loss: 0.3245556652545929
step: 150, loss: 0.2816980183124542
step: 160, loss: 0.22726327180862427
step: 170, loss: 0.3805049657821655
step: 180, loss: 0.2776198089122772
step: 190, loss: 0.11761442571878433
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.5224719101123596, f1=0.5552560646900269,
best_f1=0.5552560646900269
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24778851866722107
step: 10, loss: 0.026824988424777985
step: 20, loss: 0.09230923652648926
step: 30, loss: 0.11862214654684067
step: 40, loss: 0.35398489236831665
step: 50, loss: 0.3025135397911072
step: 60, loss: 0.05446271970868111
step: 70, loss: 0.2861699163913727
step: 80, loss: 0.07828519493341446
step: 90, loss: 0.1419447511434555
step: 100, loss: 0.23656366765499115
step: 110, loss: 0.2734087109565735
step: 120, loss: 0.15770751237869263
step: 130, loss: 0.1801139861345291
step: 140, loss: 0.20871688425540924
step: 150, loss: 0.08190395683050156
step: 160, loss: 0.07135500013828278
step: 170, loss: 0.15009044110774994
step: 180, loss: 0.13555899262428284

```

```

step: 190, loss: 0.18781065940856934
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.7202072538860104, f1=0.7545219638242895,
best_f1=0.7545219638242895
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1422896683216095
step: 10, loss: 0.16565091907978058
step: 20, loss: 0.08732151240110397
step: 30, loss: 0.047972556203603745
step: 40, loss: 0.07071816921234131
step: 50, loss: 0.21093672513961792
step: 60, loss: 0.12507843971252441
step: 70, loss: 0.09375147521495819
step: 80, loss: 0.29355689883232117
step: 90, loss: 0.07204095274209976
step: 100, loss: 0.06711848825216293
step: 110, loss: 0.21066242456436157
step: 120, loss: 0.02179543487727642
step: 130, loss: 0.05502144619822502
step: 140, loss: 0.07113868743181229
step: 150, loss: 0.04368450120091438
step: 160, loss: 0.19579589366912842
step: 170, loss: 0.03118308261036873
step: 180, loss: 0.027273626998066902
step: 190, loss: 0.06270450353622437
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.7076923076923077, f1=0.7563451776649746,
best_f1=0.7545219638242895
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09174656867980957
step: 10, loss: 0.05724283680319786
step: 20, loss: 0.01078318152576685
step: 30, loss: 0.036733947694301605
step: 40, loss: 0.0016461920458823442
step: 50, loss: 0.0459882877767086
step: 60, loss: 0.06845702230930328

```



```

step: 70, loss: 0.053839728236198425
step: 80, loss: 0.1091696098446846
step: 90, loss: 0.038241803646087646
step: 100, loss: 0.032656531780958176
step: 110, loss: 0.022449687123298645
step: 120, loss: 0.035889290273189545
step: 130, loss: 0.22849100828170776
step: 140, loss: 0.06306694447994232
step: 150, loss: 0.009505058638751507
step: 160, loss: 0.013948112726211548
step: 170, loss: 0.014963226392865181
step: 180, loss: 0.11380577087402344
step: 190, loss: 0.09618840366601944
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.7299270072992701, f1=0.7469879518072289,
best_f1=0.7469879518072289
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.059454478323459625
step: 10, loss: 0.010481752455234528
step: 20, loss: 0.10073141753673553
step: 30, loss: 0.020368952304124832
step: 40, loss: 0.031406015157699585
step: 50, loss: 0.04925308749079704
step: 60, loss: 0.015189778059720993
step: 70, loss: 0.015068071894347668
step: 80, loss: 0.007868095301091671
step: 90, loss: 0.013168470934033394
step: 100, loss: 0.08731553703546524
step: 110, loss: 0.005712681915611029
step: 120, loss: 0.0010513000888749957
step: 130, loss: 0.08832273632287979
step: 140, loss: 0.00578511506319046
step: 150, loss: 0.013316965661942959
step: 160, loss: 0.017083970829844475
step: 170, loss: 0.03397693485021591
step: 180, loss: 0.014972754754126072
step: 190, loss: 0.12271370738744736
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

epoch 5: dev_f1=0.7252124645892353, f1=0.7774647887323942,
best_f1=0.7469879518072289
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0049972571432590485
step: 10, loss: 0.0018856661627069116
step: 20, loss: 0.0016353005776181817
step: 30, loss: 0.02315210923552513
step: 40, loss: 0.040767598897218704
step: 50, loss: 0.13983412086963654
step: 60, loss: 0.010604835115373135
step: 70, loss: 0.11851274222135544
step: 80, loss: 0.2790054976940155
step: 90, loss: 0.062453605234622955
step: 100, loss: 0.030561743304133415
step: 110, loss: 0.043512310832738876
step: 120, loss: 0.016711736097931862
step: 130, loss: 0.016012031584978104
step: 140, loss: 0.02317562885582447
step: 150, loss: 0.02465434931218624
step: 160, loss: 0.01455833949148655
step: 170, loss: 0.002930481918156147
step: 180, loss: 0.006165823899209499
step: 190, loss: 0.11878831684589386
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.7022471910112359, f1=0.7548209366391184,
best_f1=0.7469879518072289
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004896987229585648
step: 10, loss: 0.0036544816102832556
step: 20, loss: 0.0018422385910525918
step: 30, loss: 0.005664553493261337
step: 40, loss: 0.008792746812105179
step: 50, loss: 0.03959572687745094
step: 60, loss: 0.0037444184999912977
step: 70, loss: 0.0027255266904830933
step: 80, loss: 0.006766860838979483
step: 90, loss: 0.11486638337373734
step: 100, loss: 0.003186532761901617
step: 110, loss: 0.008279488421976566
step: 120, loss: 0.007562992163002491
step: 130, loss: 0.0028132302686572075
step: 140, loss: 0.004265196621417999

```

```

step: 150, loss: 0.06513139605522156
step: 160, loss: 0.023053184151649475
step: 170, loss: 0.0030198292806744576
step: 180, loss: 0.057274237275123596
step: 190, loss: 0.004212684463709593
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.7142857142857142, f1=0.7492795389048991,
best_f1=0.7469879518072289
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007795641198754311
step: 10, loss: 0.003935668617486954
step: 20, loss: 0.010378717444837093
step: 30, loss: 0.002071820665150881
step: 40, loss: 0.001493839779868722
step: 50, loss: 0.0009302273974753916
step: 60, loss: 0.12776684761047363
step: 70, loss: 0.011559219099581242
step: 80, loss: 0.05390920117497444
step: 90, loss: 0.020068570971488953
step: 100, loss: 0.040294941514730453
step: 110, loss: 0.005619569215923548
step: 120, loss: 0.0012854719534516335
step: 130, loss: 0.0027761796955019236
step: 140, loss: 0.0016558386851102114
step: 150, loss: 0.0025828275829553604
step: 160, loss: 0.000943192804697901
step: 170, loss: 0.0010017899330705404
step: 180, loss: 0.0011448945151641965
step: 190, loss: 0.03843311965465546
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.7150537634408601, f1=0.7598944591029023,
best_f1=0.7469879518072289
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.030142877250909805
step: 10, loss: 0.023142913356423378
step: 20, loss: 0.02949501760303974
step: 30, loss: 0.08340156078338623
step: 40, loss: 0.012426825240254402
step: 50, loss: 0.0010370189556851983
step: 60, loss: 0.003057509195059538

```

```

step: 70, loss: 0.0017089274479076266
step: 80, loss: 0.026993384584784508
step: 90, loss: 0.02218233421444893
step: 100, loss: 0.002163098892197013
step: 110, loss: 0.006055066362023354
step: 120, loss: 0.017252124845981598
step: 130, loss: 0.001735617290250957
step: 140, loss: 0.0022797260899096727
step: 150, loss: 0.021944042295217514
step: 160, loss: 0.18360863626003265
step: 170, loss: 0.012550114654004574
step: 180, loss: 0.028573554009199142
step: 190, loss: 0.005345471668988466
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.7226890756302521, f1=0.7507002801120448,
best_f1=0.7469879518072289
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000630514114163816
step: 10, loss: 0.08352042734622955
step: 20, loss: 0.0013707653852179646
step: 30, loss: 0.004141259007155895
step: 40, loss: 0.0018456332618370652
step: 50, loss: 0.028775524348020554
step: 60, loss: 0.021223824471235275
step: 70, loss: 0.014237258583307266
step: 80, loss: 0.003618821734562516
step: 90, loss: 0.0059831817634403706
step: 100, loss: 0.06708399951457977
step: 110, loss: 0.057613059878349304
step: 120, loss: 0.019549861550331116
step: 130, loss: 0.09484650939702988
step: 140, loss: 0.007794151082634926
step: 150, loss: 0.0011618067510426044
step: 160, loss: 0.0029460524674504995
step: 170, loss: 0.0004696861724369228
step: 180, loss: 0.0008319318876601756
step: 190, loss: 0.06844649463891983
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.7258064516129032, f1=0.783289817232376,
best_f1=0.7469879518072289
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005477493978105485
step: 10, loss: 0.004931125324219465
step: 20, loss: 0.0011360968928784132
step: 30, loss: 0.00774182565510273
step: 40, loss: 0.04312427341938019
step: 50, loss: 0.0016045207157731056
step: 60, loss: 0.0006883378373458982
step: 70, loss: 0.000934453506488353
step: 80, loss: 0.0033390328753739595
step: 90, loss: 0.02239808812737465
step: 100, loss: 0.04242763668298721
step: 110, loss: 0.00034914346178993583
step: 120, loss: 0.0037384317256510258
step: 130, loss: 0.0013438249006867409
step: 140, loss: 0.00046054559061303735
step: 150, loss: 0.0004343200125731528
step: 160, loss: 0.0019341135630384088
step: 170, loss: 0.0038640745915472507
step: 180, loss: 0.035741012543439865
step: 190, loss: 0.005867965053766966
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.707774798927614, f1=0.7708894878706198,
best_f1=0.7469879518072289
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013620456447824836
step: 10, loss: 0.003782485844567418
step: 20, loss: 0.002003952395170927
step: 30, loss: 0.15166695415973663
step: 40, loss: 0.00309711298905313
step: 50, loss: 0.017287477850914
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.0004009094263892621
step: 70, loss: 0.0061729783192276955
step: 80, loss: 0.0008630852098576725
step: 90, loss: 0.0006517742876894772
step: 100, loss: 0.001917241490446031
step: 110, loss: 0.005338811781257391
step: 120, loss: 0.00044286029879003763
step: 130, loss: 0.0009752405458129942
step: 140, loss: 0.05417175590991974
step: 150, loss: 0.0008786192047409713
step: 160, loss: 0.003513244679197669
step: 170, loss: 0.00878090225160122

```

```

step: 180, loss: 0.0007612554472871125
step: 190, loss: 0.027628760784864426
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.7175141242937854, f1=0.7648725212464589,
best_f1=0.7469879518072289
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017360621131956577
step: 10, loss: 0.005967312958091497
step: 20, loss: 0.003937283530831337
step: 30, loss: 0.004087241366505623
step: 40, loss: 0.0007932699518278241
step: 50, loss: 0.01608286052942276
step: 60, loss: 0.0006606621900573373
step: 70, loss: 0.0007182210683822632
step: 80, loss: 0.002742298413068056
step: 90, loss: 0.015053441748023033
step: 100, loss: 0.059965699911117554
step: 110, loss: 0.0008541177376173437
step: 120, loss: 0.0030367691069841385
step: 130, loss: 0.0010088077979162335
step: 140, loss: 0.02664179354906082
step: 150, loss: 0.002947412431240082
step: 160, loss: 0.0008971297065727413
step: 170, loss: 0.00043364582234062254
step: 180, loss: 0.011564998887479305
step: 190, loss: 0.0004701162106357515
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.7068493150684931, f1=0.7811634349030471,
best_f1=0.7469879518072289
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012172162532806396
step: 10, loss: 0.0006760658579878509
step: 20, loss: 0.0005041694967076182
step: 30, loss: 0.005573469679802656
step: 40, loss: 0.003882349468767643
step: 50, loss: 0.02182987704873085
step: 60, loss: 0.0504813976585865
step: 70, loss: 0.0030211801640689373
step: 80, loss: 0.0017080778488889337
step: 90, loss: 0.0037923436611890793

```

```

step: 100, loss: 0.0017576093086972833
step: 110, loss: 0.0015738771762698889
step: 120, loss: 0.004429524298757315
step: 130, loss: 0.020857304334640503
step: 140, loss: 0.011322715319693089
step: 150, loss: 0.0006882477318868041
step: 160, loss: 0.11673617362976074
step: 170, loss: 0.0012810979969799519
step: 180, loss: 0.0057492610067129135
step: 190, loss: 0.0004149639862589538
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.705539358600583, f1=0.7441860465116279,
best_f1=0.7469879518072289
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04937971755862236
step: 10, loss: 0.0035852829460054636
step: 20, loss: 0.0013250623596832156
step: 30, loss: 0.0010083138477057219
step: 40, loss: 0.0008713380084373057
step: 50, loss: 0.011477962136268616
step: 60, loss: 0.0009741755202412605
step: 70, loss: 0.03294241800904274
step: 80, loss: 0.010288123972713947
step: 90, loss: 0.02413492277264595
step: 100, loss: 0.0036534410901367664
step: 110, loss: 0.005110010039061308
step: 120, loss: 0.0008959955885075033
step: 130, loss: 0.0016364236362278461
step: 140, loss: 0.17377524077892303
step: 150, loss: 0.016101734712719917
step: 160, loss: 0.0038997260853648186
step: 170, loss: 0.013128484599292278
step: 180, loss: 0.001026240410283208
step: 190, loss: 0.024325529113411903
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.7058823529411764, f1=0.7470588235294117,
best_f1=0.7469879518072289
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:05, 354.84it/s]
load_f1 = 0.6235011990407673
real_f1 = 0.6117647058823529
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:10, 426.20it/s]

```

## 4.2 DITTO TEXTUAL

### 4.2.1 Abt-Buy - Running the matcher

```

[ ]: # Step 1:
      # train the wdc model with the save_model flag on
      # it will produce two files *_dev.pt and *_test.pt which are the
      # best checkpoints on the dev set and the test set.
      !CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
        --task Textual/Abt-Buy \
        --batch_size 32 \
        --max_len 256 \
        --lr 3e-5 \

```



```

--n_epochs 15 \
--finetuning \
--lm distilbert \
--fp16 \
--save_model \
--dk general \
--dk product \
--summarize \
--da entry_swap \
--da attr_del

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Abt-Buy.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Textual/Abt-Buy \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm distilbert \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback.

```

Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.8392401337623596
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.22173160314559937
step: 20, loss: 0.15833590924739838
step: 30, loss: 0.23458170890808105
step: 40, loss: 0.3070785403251648
step: 50, loss: 0.3746965825557709
step: 60, loss: 0.448626846075058
step: 70, loss: 0.3097025752067566
step: 80, loss: 0.24999070167541504
step: 90, loss: 0.40694889426231384
step: 100, loss: 0.22261394560337067
step: 110, loss: 0.17666316032409668
step: 120, loss: 0.5441346168518066
step: 130, loss: 0.4143960773944855
step: 140, loss: 0.45193928480148315
step: 150, loss: 0.11200741678476334
step: 160, loss: 0.297050416469574
step: 170, loss: 0.1634754240512848
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.6084656084656085, f1=0.6046511627906976,
best_f1=0.6046511627906976
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.36583006381988525
step: 10, loss: 0.1735670566558838
step: 20, loss: 0.20943652093410492

```

```

step: 30, loss: 0.3031376600265503
step: 40, loss: 0.14736124873161316
step: 50, loss: 0.22644011676311493
step: 60, loss: 0.09363417327404022
step: 70, loss: 0.2667844295501709
step: 80, loss: 0.10500527918338776
step: 90, loss: 0.16993668675422668
step: 100, loss: 0.14346736669540405
step: 110, loss: 0.1664213240146637
step: 120, loss: 0.04688489809632301
step: 130, loss: 0.07307519018650055
step: 140, loss: 0.14046287536621094
step: 150, loss: 0.12859325110912323
step: 160, loss: 0.11325109750032425
step: 170, loss: 0.32557225227355957
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.7539267015706808, f1=0.7568922305764412,
best_f1=0.7568922305764412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04440565034747124
step: 10, loss: 0.05048041790723801
step: 20, loss: 0.03624534234404564
step: 30, loss: 0.17252826690673828
step: 40, loss: 0.007439074572175741
step: 50, loss: 0.12192723900079727
step: 60, loss: 0.07701601833105087
step: 70, loss: 0.10828421264886856
step: 80, loss: 0.13560059666633606
step: 90, loss: 0.13479283452033997
step: 100, loss: 0.029715023934841156
step: 110, loss: 0.13408713042736053
step: 120, loss: 0.005101468414068222
step: 130, loss: 0.11894823610782623
step: 140, loss: 0.03903589025139809
step: 150, loss: 0.022524122148752213
step: 160, loss: 0.08839389681816101
step: 170, loss: 0.08871358633041382
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

epoch 3: dev_f1=0.7338129496402879, f1=0.7703349282296651,
best_f1=0.7568922305764412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.092044398188591
step: 10, loss: 0.1572227030992508
step: 20, loss: 0.030738620087504387
step: 30, loss: 0.02390335500240326
step: 40, loss: 0.017888354137539864
step: 50, loss: 0.006528157740831375
step: 60, loss: 0.10129912197589874
step: 70, loss: 0.022390782833099365
step: 80, loss: 0.1264304369688034
step: 90, loss: 0.017896577715873718
step: 100, loss: 0.04321185126900673
step: 110, loss: 0.12090963870286942
step: 120, loss: 0.19927304983139038
step: 130, loss: 0.07222714275121689
step: 140, loss: 0.023874342441558838
step: 150, loss: 0.0790957659482956
step: 160, loss: 0.06969702243804932
step: 170, loss: 0.019426722079515457
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.7518796992481204, f1=0.7737226277372263,
best_f1=0.7568922305764412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.044411566108465195
step: 10, loss: 0.026743248105049133
step: 20, loss: 0.040189456194639206
step: 30, loss: 0.06808511912822723
step: 40, loss: 0.009450940415263176
step: 50, loss: 0.043602354824543
step: 60, loss: 0.013012652285397053
step: 70, loss: 0.02643812820315361
step: 80, loss: 0.0139684546738863
step: 90, loss: 0.026336219161748886
step: 100, loss: 0.02541303262114525
step: 110, loss: 0.023236030712723732
step: 120, loss: 0.04176778718829155
step: 130, loss: 0.10472523421049118
step: 140, loss: 0.08415985852479935
step: 150, loss: 0.017181698232889175
step: 160, loss: 0.024336738511919975

```

```

step: 170, loss: 0.0611811988055706
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.7532467532467533, f1=0.792079207920792,
best_f1=0.7568922305764412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016584327444434166
step: 10, loss: 0.07944797724485397
step: 20, loss: 0.019664959982037544
step: 30, loss: 0.032585062086582184
step: 40, loss: 0.013022211380302906
step: 50, loss: 0.21790724992752075
step: 60, loss: 0.028464462608098984
step: 70, loss: 0.06731788069009781
step: 80, loss: 0.18604187667369843
step: 90, loss: 0.05078698694705963
step: 100, loss: 0.011149383150041103
step: 110, loss: 0.010385953821241856
step: 120, loss: 0.044777266681194305
step: 130, loss: 0.19631721079349518
step: 140, loss: 0.021315909922122955
step: 150, loss: 0.24968989193439484
step: 160, loss: 0.07441265881061554
step: 170, loss: 0.010588949546217918
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 6: dev_f1=0.7548076923076923, f1=0.7772727272727272,
best_f1=0.7772727272727272
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010654640384018421
step: 10, loss: 0.0009637334733270109
step: 20, loss: 0.0032107126899063587
step: 30, loss: 0.08828329294919968
step: 40, loss: 0.10638436675071716
step: 50, loss: 0.014643503352999687
step: 60, loss: 0.27112331986427307
step: 70, loss: 0.02387259714305401
step: 80, loss: 0.0126262828707695

```

```

step: 90, loss: 0.007621956989169121
step: 100, loss: 0.047642067074775696
step: 110, loss: 0.0016893886495381594
step: 120, loss: 0.16963298618793488
step: 130, loss: 0.2559235692024231
step: 140, loss: 0.1153542771935463
step: 150, loss: 0.07261596620082855
step: 160, loss: 0.03004707396030426
step: 170, loss: 0.1527077853679657
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.7350835322195706, f1=0.744920993227991,
best_f1=0.7772727272727272
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020747140049934387
step: 10, loss: 0.05342225357890129
step: 20, loss: 0.007359989918768406
step: 30, loss: 0.03456268459558487
step: 40, loss: 0.005135382059961557
step: 50, loss: 0.04763489216566086
step: 60, loss: 0.006109511014074087
step: 70, loss: 0.02957480028271675
step: 80, loss: 0.003996869083493948
step: 90, loss: 0.05959802493453026
step: 100, loss: 0.039968062192201614
step: 110, loss: 0.018115125596523285
step: 120, loss: 0.001188627677038312
step: 130, loss: 0.03511965647339821
step: 140, loss: 0.08615873754024506
step: 150, loss: 0.19454646110534668
step: 160, loss: 0.03162779659032822
step: 170, loss: 0.007022487465292215
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 8: dev_f1=0.7579462102689486, f1=0.7933491686460807,
best_f1=0.7933491686460807
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02149989642202854

```

```

step: 10, loss: 0.0093683497980237
step: 20, loss: 0.058717239648103714
step: 30, loss: 0.013009625487029552
step: 40, loss: 0.02831832505762577
step: 50, loss: 0.016464300453662872
step: 60, loss: 0.004725801292806864
step: 70, loss: 0.022418532520532608
step: 80, loss: 0.09127926081418991
step: 90, loss: 0.025644592940807343
step: 100, loss: 0.1047392413020134
step: 110, loss: 0.0017886589048430324
step: 120, loss: 0.005185144487768412
step: 130, loss: 0.003011928638443351
step: 140, loss: 0.0016215380746871233
step: 150, loss: 0.006378337740898132
step: 160, loss: 0.019106173887848854
step: 170, loss: 0.02733554132282734
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.7486631016042781, f1=0.7794871794871795,
best_f1=0.7933491686460807
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0036379308439791203
step: 10, loss: 0.056188009679317474
step: 20, loss: 0.0029223000165075064
step: 30, loss: 0.0011401617666706443
step: 40, loss: 0.08647676557302475
step: 50, loss: 0.013422493822872639
step: 60, loss: 0.02367507666349411
step: 70, loss: 0.019475774839520454
step: 80, loss: 0.011200402863323689
step: 90, loss: 0.049785126000642776
step: 100, loss: 0.020616263151168823
step: 110, loss: 0.06958489120006561
step: 120, loss: 0.029627133160829544
step: 130, loss: 0.10426245629787445
step: 140, loss: 0.023479141294956207
step: 150, loss: 0.010075950063765049
step: 160, loss: 0.014738176017999649
step: 170, loss: 0.026370784267783165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or

```

```

loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 10: dev_f1=0.7598039215686274, f1=0.782816229116945,
best_f1=0.782816229116945
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005577914998866618
step: 10, loss: 0.012685786932706833
step: 20, loss: 0.01465071365237236
step: 30, loss: 0.0010318058775737882
step: 40, loss: 0.007979197427630424
step: 50, loss: 0.016119061037898064
step: 60, loss: 0.004913675133138895
step: 70, loss: 0.0015756472712382674
step: 80, loss: 0.0012590551050379872
step: 90, loss: 0.0012639597989618778
step: 100, loss: 0.03479408845305443
step: 110, loss: 0.009279552847146988
step: 120, loss: 0.07630547881126404
step: 130, loss: 0.001989514334127307
step: 140, loss: 0.039398860186338425
step: 150, loss: 0.004981605801731348
step: 160, loss: 0.0035852708388119936
step: 170, loss: 0.13122226297855377
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.7389162561576353, f1=0.7971360381861575,
best_f1=0.782816229116945
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012994328513741493
step: 10, loss: 0.10009069740772247
step: 20, loss: 0.0595097541809082
step: 30, loss: 0.0026952195912599564
step: 40, loss: 0.006947800517082214
step: 50, loss: 0.011787306517362595
step: 60, loss: 0.012716429308056831
step: 70, loss: 0.04824940860271454
step: 80, loss: 0.006288356613367796
step: 90, loss: 0.0020516221411526203
step: 100, loss: 0.011528817005455494
step: 110, loss: 0.0006694855401292443
step: 120, loss: 0.007812227588146925
step: 130, loss: 0.007188885938376188
step: 140, loss: 0.004477239213883877

```



```

step: 150, loss: 0.04581499844789505
step: 160, loss: 0.0007757695857435465
step: 170, loss: 0.002638885984197259
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.7512437810945273, f1=0.7777777777777778,
best_f1=0.782816229116945
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021932919044047594
step: 10, loss: 0.07728724181652069
step: 20, loss: 0.00690992409363389
step: 30, loss: 0.0025037515442818403
step: 40, loss: 0.00847319234162569
step: 50, loss: 0.0030435826629400253
step: 60, loss: 0.010078277438879013
step: 70, loss: 0.008833348751068115
step: 80, loss: 0.0013094600290060043
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 90, loss: 0.0013474458828568459
step: 100, loss: 0.0013466469245031476
step: 110, loss: 0.002735004061833024
step: 120, loss: 0.0011155783431604505
step: 130, loss: 0.017196115106344223
step: 140, loss: 0.0005073765642009676
step: 150, loss: 0.0005625606863759458
step: 160, loss: 0.006169267930090427
step: 170, loss: 0.003649403341114521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 13: dev_f1=0.760204081632653, f1=0.7990074441687345,
best_f1=0.7990074441687345
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028073431458324194
step: 10, loss: 0.07043298333883286
step: 20, loss: 0.07453517615795135
step: 30, loss: 0.003768552327528596
step: 40, loss: 0.0013288829941302538
step: 50, loss: 0.24349166452884674

```

```

step: 60, loss: 0.003977038431912661
step: 70, loss: 0.017816418781876564
step: 80, loss: 0.0017193254316225648
step: 90, loss: 0.00628051208332181
step: 100, loss: 0.0006073163822293282
step: 110, loss: 0.019129883497953415
step: 120, loss: 0.011152335442602634
step: 130, loss: 0.00240717688575387
step: 140, loss: 0.006490116007626057
step: 150, loss: 0.0005268146633170545
step: 160, loss: 0.03775409981608391
step: 170, loss: 0.0020661999005824327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.7566265060240964, f1=0.7981438515081206,
best_f1=0.7990074441687345
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03822333738207817
step: 10, loss: 0.012777047231793404
step: 20, loss: 0.05472162738442421
step: 30, loss: 0.007555073592811823
step: 40, loss: 0.04878167808055878
step: 50, loss: 0.0033698093611747026
step: 60, loss: 0.001267191837541759
step: 70, loss: 0.01565287634730339
step: 80, loss: 0.014921103604137897
step: 90, loss: 0.0003243952232878655
step: 100, loss: 0.002375779440626502
step: 110, loss: 0.00040535975131206214
step: 120, loss: 0.037441570311784744
step: 130, loss: 0.0022286272142082453
step: 140, loss: 0.008809665217995644
step: 150, loss: 0.05694270133972168
step: 160, loss: 0.00034445011988282204
step: 170, loss: 0.005075869150459766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.7572815533980582, f1=0.7962529274004684,
best_f1=0.7990074441687345
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
1916it [00:04, 462.70it/s]
load_f1 = 0.5683836589698046
real_f1 = 0.5519591141396933
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:10, 414.89it/s]

```

## 4.3 DITTO DIRTY

### 4.3.1 DBLP-ACM - Running the matcher

```

[ ]: # Step 1:
      # train the wdc model with the save_model flag on
      # it will produce two files *_dev.pt and *_test.pt which are the
      # best checkpoints on the dev set and the test set.
      !CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
        --task Dirty/DBLP-ACM \
        --batch_size 32 \
        --max_len 256 \
        --lr 3e-5 \

```

```

--n_epochs 15 \
--finetuning \
--lm distilbert \
--fp16 \
--save_model \
--dk general \
--dk product \
--summarize \
--da entry_swap \
--da attr_del

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-ACM.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Dirty/DBLP-ACM \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm distilbert \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback.

```

Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.7915582060813904
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4780407249927521
step: 20, loss: 0.632552981376648
step: 30, loss: 0.4918677508831024
step: 40, loss: 0.3381776511669159
step: 50, loss: 0.14120005071163177
step: 60, loss: 0.13378234207630157
step: 70, loss: 0.23915068805217743
step: 80, loss: 0.2506563067436218
step: 90, loss: 0.04648857191205025
step: 100, loss: 0.1368880569934845
step: 110, loss: 0.11390628665685654
step: 120, loss: 0.05225758254528046
step: 130, loss: 0.022069675847887993
step: 140, loss: 0.03528766706585884
step: 150, loss: 0.15808509290218353
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 160, loss: 0.25333738327026367
step: 170, loss: 0.02349790558218956
step: 180, loss: 0.011617986485362053
step: 190, loss: 0.03656519949436188
step: 200, loss: 0.0771552175283432
step: 210, loss: 0.027849307283759117
step: 220, loss: 0.03356248512864113
step: 230, loss: 0.023096993565559387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.950561797752809, f1=0.9548532731376976,

```

```

best_f1=0.9548532731376976
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.054700784385204315
step: 10, loss: 0.06914638727903366
step: 20, loss: 0.004425074905157089
step: 30, loss: 0.00947518553584814
step: 40, loss: 0.05514531582593918
step: 50, loss: 0.020694902166724205
step: 60, loss: 0.009628959000110626
step: 70, loss: 0.017306044697761536
step: 80, loss: 0.01784011907875538
step: 90, loss: 0.015182366594672203
step: 100, loss: 0.10424626618623734
step: 110, loss: 0.11272326856851578
step: 120, loss: 0.07670973241329193
step: 130, loss: 0.00818928238004446
step: 140, loss: 0.39255306124687195
step: 150, loss: 0.11405043303966522
step: 160, loss: 0.01985057070851326
step: 170, loss: 0.0934809148311615
step: 180, loss: 0.023226335644721985
step: 190, loss: 0.1447497308254242
step: 200, loss: 0.006399459671229124
step: 210, loss: 0.04325556382536888
step: 220, loss: 0.001810698420740664
step: 230, loss: 0.021960893645882607
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9623059866962307, f1=0.9620535714285715,
best_f1=0.9620535714285715
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11526811122894287
step: 10, loss: 0.0741630271077156
step: 20, loss: 0.054142214357852936
step: 30, loss: 0.04603252559900284
step: 40, loss: 0.06120741739869118
step: 50, loss: 0.005759149789810181
step: 60, loss: 0.05421421304345131
step: 70, loss: 0.0027882009744644165

```

```

step: 80, loss: 0.0009462390444241464
step: 90, loss: 0.18099310994148254
step: 100, loss: 0.11837457865476608
step: 110, loss: 0.08258862793445587
step: 120, loss: 0.06044322997331619
step: 130, loss: 0.005078524816781282
step: 140, loss: 0.004412806127220392
step: 150, loss: 0.002851656172424555
step: 160, loss: 0.022579286247491837
step: 170, loss: 0.016115380451083183
step: 180, loss: 0.004053664393723011
step: 190, loss: 0.01834096573293209
step: 200, loss: 0.005272995214909315
step: 210, loss: 0.01911275088787079
step: 220, loss: 0.21025167405605316
step: 230, loss: 0.17289866507053375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.9597315436241611, f1=0.9589345172031077,
best_f1=0.9620535714285715
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007739463355392218
step: 10, loss: 0.024776823818683624
step: 20, loss: 0.0022280991543084383
step: 30, loss: 0.015351939015090466
step: 40, loss: 0.01862122118473053
step: 50, loss: 0.005774314980953932
step: 60, loss: 0.001091106329113245
step: 70, loss: 0.09202464669942856
step: 80, loss: 0.18335743248462677
step: 90, loss: 0.03739995136857033
step: 100, loss: 0.0026919792871922255
step: 110, loss: 0.006888791918754578
step: 120, loss: 0.07044021785259247
step: 130, loss: 0.018132170662283897
step: 140, loss: 0.003511902177706361
step: 150, loss: 0.008401118218898773
step: 160, loss: 0.02136332355439663
step: 170, loss: 0.006759267766028643
step: 180, loss: 0.04051962494850159
step: 190, loss: 0.006787593942135572
step: 200, loss: 0.01685423031449318
step: 210, loss: 0.0005536031094379723
step: 220, loss: 0.0021651138085871935
step: 230, loss: 0.0005791114526800811

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.9608938547486034, f1=0.9620535714285715,
best_f1=0.9620535714285715
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007128422148525715
step: 10, loss: 0.004270216915756464
step: 20, loss: 0.001005941885523498
step: 30, loss: 0.0006610004929825664
step: 40, loss: 0.00531422533094883
step: 50, loss: 0.0030871655326336622
step: 60, loss: 0.018251921981573105
step: 70, loss: 0.033237967640161514
step: 80, loss: 0.0032550375908613205
step: 90, loss: 0.05770629644393921
step: 100, loss: 0.05176214873790741
step: 110, loss: 0.011815489269793034
step: 120, loss: 0.0644628256559372
step: 130, loss: 0.10144862532615662
step: 140, loss: 0.0006138848839327693
step: 150, loss: 0.00023500066890846938
step: 160, loss: 0.02989777736365795
step: 170, loss: 0.2990092635154724
step: 180, loss: 0.004199079237878323
step: 190, loss: 0.00707911467179656
step: 200, loss: 0.0022141337394714355
step: 210, loss: 0.002066215267404914
step: 220, loss: 0.00038806229713372886
step: 230, loss: 0.0017322488129138947
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.9688888888888889, f1=0.9666666666666666,
best_f1=0.9666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01105063408613205
step: 10, loss: 0.027630051597952843
step: 20, loss: 0.0027562060859054327
step: 30, loss: 0.0010495878523215652

```



```

step: 40, loss: 0.0019289656775072217
step: 50, loss: 0.12998811900615692
step: 60, loss: 0.018982447683811188
step: 70, loss: 0.001661514164879918
step: 80, loss: 0.0019930810667574406
step: 90, loss: 0.000796738953795284
step: 100, loss: 0.000537979940418154
step: 110, loss: 0.05554182454943657
step: 120, loss: 0.0004246170283295214
step: 130, loss: 0.007205284666270018
step: 140, loss: 0.0006684178370051086
step: 150, loss: 0.0068577369675040245
step: 160, loss: 0.00573695870116353
step: 170, loss: 0.0003867620835080743
step: 180, loss: 0.00039250258123502135
step: 190, loss: 0.10987986624240875
step: 200, loss: 0.0814061164855957
step: 210, loss: 0.028824210166931152
step: 220, loss: 0.0016058428445830941
step: 230, loss: 0.0005519058904610574
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9684684684684683, f1=0.9663677130044843,
best_f1=0.9666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08834263682365417
step: 10, loss: 0.005911382380872965
step: 20, loss: 0.00038169726030901074
step: 30, loss: 0.00041079241782426834
step: 40, loss: 0.010633165016770363
step: 50, loss: 0.00019306526519358158
step: 60, loss: 0.003275267779827118
step: 70, loss: 0.0006792506319470704
step: 80, loss: 8.759603224461898e-05
step: 90, loss: 0.020341377705335617
step: 100, loss: 0.004817801993340254
step: 110, loss: 0.0014620316214859486
step: 120, loss: 0.01716344617307186
step: 130, loss: 0.00382046215236187
step: 140, loss: 0.00011119140981463715
step: 150, loss: 0.0020802973303943872
step: 160, loss: 0.0004347854992374778
step: 170, loss: 9.486118506174535e-05
step: 180, loss: 0.00032147939782589674
step: 190, loss: 0.00011155161337228492

```

```

step: 200, loss: 0.006077934987843037
step: 210, loss: 0.0029670805670320988
step: 220, loss: 0.008737044408917427
step: 230, loss: 0.03365939110517502
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9584736251402918, f1=0.9631284916201117,
best_f1=0.9666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03446396440267563
step: 10, loss: 0.024299975484609604
step: 20, loss: 0.0016066894168034196
step: 30, loss: 0.001974348444491625
step: 40, loss: 0.0054787578992545605
step: 50, loss: 0.003653265768662095
step: 60, loss: 0.0005837006028741598
step: 70, loss: 0.0002702598867472261
step: 80, loss: 0.0017781957285478711
step: 90, loss: 0.00044967798748984933
step: 100, loss: 0.00017962348647415638
step: 110, loss: 0.0009041334269568324
step: 120, loss: 0.00011747271491913125
step: 130, loss: 0.007637544069439173
step: 140, loss: 0.0015165569493547082
step: 150, loss: 0.005048295482993126
step: 160, loss: 0.002764600794762373
step: 170, loss: 0.00038142898119986057
step: 180, loss: 0.0008652090909898083
step: 190, loss: 0.0003144355723634362
step: 200, loss: 0.004605717025697231
step: 210, loss: 0.02828916348516941
step: 220, loss: 0.0010467271786183119
step: 230, loss: 0.06600001454353333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9662921348314607, f1=0.9632107023411371,
best_f1=0.9666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004374620970338583
step: 10, loss: 0.0060041239485144615
step: 20, loss: 0.0004102336533833295
step: 30, loss: 0.00037969963159412146

```

```

step: 40, loss: 5.085706288809888e-05
step: 50, loss: 0.0007281400612555444
step: 60, loss: 0.0024545961059629917
step: 70, loss: 0.06020532175898552
step: 80, loss: 0.051985424011945724
step: 90, loss: 0.0016614010091871023
step: 100, loss: 0.003903749631717801
step: 110, loss: 0.0007006602245382965
step: 120, loss: 0.026153864338994026
step: 130, loss: 0.0001090119912987575
step: 140, loss: 0.001136177801527083
step: 150, loss: 8.60266518429853e-05
step: 160, loss: 0.000424231548095122
step: 170, loss: 0.0026243487372994423
step: 180, loss: 0.000297749531455338
step: 190, loss: 0.00021817559900227934
step: 200, loss: 0.001219955855049193
step: 210, loss: 0.0003080317110288888
step: 220, loss: 0.022622125223279
step: 230, loss: 0.09290561825037003
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.970917225950783, f1=0.9665924276169264,
best_f1=0.9665924276169264
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00023566292657051235
step: 10, loss: 0.00045520366984419525
step: 20, loss: 0.1501949280500412
step: 30, loss: 0.00029438722413033247
step: 40, loss: 0.0005756454775109887
step: 50, loss: 0.018866194412112236
step: 60, loss: 0.02852984145283699
step: 70, loss: 0.0006082315230742097
step: 80, loss: 0.0025794191751629114
step: 90, loss: 0.0050967177376151085
step: 100, loss: 0.00020058271184097975
step: 110, loss: 0.0004913138691335917
step: 120, loss: 0.02207597717642784
step: 130, loss: 0.003112158738076687
step: 140, loss: 0.050247207283973694
step: 150, loss: 0.0004745198821183294

```

```

step: 160, loss: 0.00018220618949271739
step: 170, loss: 0.0008319778717122972
step: 180, loss: 0.005072528030723333
step: 190, loss: 0.0020084695424884558
step: 200, loss: 7.26151920389384e-05
step: 210, loss: 0.00012164026702521369
step: 220, loss: 0.0016735476674512029
step: 230, loss: 0.0007599602686241269
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9671574178935448, f1=0.9617977528089887,
best_f1=0.9665924276169264
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.948356116889045e-05
step: 10, loss: 0.00016692040662746876
step: 20, loss: 4.803899355465546e-05
step: 30, loss: 0.0022622845135629177
step: 40, loss: 0.0012792365159839392
step: 50, loss: 0.001140741864219308
step: 60, loss: 0.0007744571776129305
step: 70, loss: 0.0016939210472628474
step: 80, loss: 0.0001232240756507963
step: 90, loss: 0.00017033099720720202
step: 100, loss: 8.411485032411292e-05
step: 110, loss: 0.00011421778617659584
step: 120, loss: 0.00020945158030372113
step: 130, loss: 5.3571875469060615e-05
step: 140, loss: 0.033108584582805634
step: 150, loss: 4.6150591515470296e-05
step: 160, loss: 0.00013454796862788498
step: 170, loss: 0.005909590050578117
step: 180, loss: 0.0017638655845075846
step: 190, loss: 0.00026908988365903497
step: 200, loss: 0.001061584334820509
step: 210, loss: 5.429311204352416e-05
step: 220, loss: 8.309996337629855e-05
step: 230, loss: 0.017057303339242935
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9694915254237287, f1=0.9628796400449944,
best_f1=0.9665924276169264
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0002044559078058228
step: 10, loss: 0.009248408488929272
step: 20, loss: 0.0002872803888749331
step: 30, loss: 0.0006002248846925795
step: 40, loss: 0.0001340572052868083
step: 50, loss: 0.00013286415196489543
step: 60, loss: 0.014048771001398563
step: 70, loss: 0.0021193602588027716
step: 80, loss: 6.341450352920219e-05
step: 90, loss: 0.00018600189650896937
step: 100, loss: 6.154591392260045e-05
step: 110, loss: 0.002655760617926717
step: 120, loss: 8.408234862145036e-05
step: 130, loss: 0.05581020936369896
step: 140, loss: 0.00016502804646734148
step: 150, loss: 0.00011607431224547327
step: 160, loss: 0.014486967585980892
step: 170, loss: 0.00016672942729201168
step: 180, loss: 0.00014838324568700045
step: 190, loss: 0.0002869763702619821
step: 200, loss: 0.02404186502099037
step: 210, loss: 0.00014259842282626778
step: 220, loss: 0.015553392469882965
step: 230, loss: 0.00010049101547338068
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 12: dev_f1=0.9742441209406495, f1=0.9620535714285715,
best_f1=0.9620535714285715
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016563957615289837
step: 10, loss: 0.000273197831120342
step: 20, loss: 8.404631080338731e-05
step: 30, loss: 0.01580977998673916
step: 40, loss: 0.00016087219410110265
step: 50, loss: 0.00018397931125946343
step: 60, loss: 0.0002165961777791381
step: 70, loss: 4.7726527554914355e-05
step: 80, loss: 0.09427032619714737
step: 90, loss: 0.0001065219403244555
step: 100, loss: 0.0001898852933663875
step: 110, loss: 0.0001027641847031191

```

```

step: 120, loss: 0.0004906412796117365
step: 130, loss: 0.00022492362768389285
step: 140, loss: 0.015510453842580318
step: 150, loss: 0.025885120034217834
step: 160, loss: 0.00019910615810658783
step: 170, loss: 8.963324216892943e-05
step: 180, loss: 0.00017185445176437497
step: 190, loss: 0.0001369460514979437
step: 200, loss: 0.00021190133702475578
step: 210, loss: 8.820098446449265e-05
step: 220, loss: 0.0005193178658373654
step: 230, loss: 0.00016963620146270841
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.972972972972973, f1=0.9664429530201343,
best_f1=0.9620535714285715
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.856304778717458e-05
step: 10, loss: 9.411947394255549e-05
step: 20, loss: 0.002180923707783222
step: 30, loss: 6.30087815807201e-05
step: 40, loss: 4.607839582604356e-05
step: 50, loss: 0.00012050052464473993
step: 60, loss: 0.018291467800736427
step: 70, loss: 0.00012038280692650005
step: 80, loss: 0.00024448009207844734
step: 90, loss: 0.00023044401314109564
step: 100, loss: 0.003494427539408207
step: 110, loss: 0.005783329717814922
step: 120, loss: 0.0001232151553267613
step: 130, loss: 0.0018109821248799562
step: 140, loss: 8.829677244648337e-05
step: 150, loss: 0.0001493136223871261
step: 160, loss: 6.064376066206023e-05
step: 170, loss: 0.00024089684302452952
step: 180, loss: 0.0001083941460819915
step: 190, loss: 0.005127369426190853
step: 200, loss: 6.663840031251311e-05
step: 210, loss: 0.023460140451788902
step: 220, loss: 0.00016236105875577778
step: 230, loss: 3.6089448258280754e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9730337078651685, f1=0.9609810479375697,

```

```

best_f1=0.9620535714285715
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003769603499677032
step: 10, loss: 0.00036492361687123775
step: 20, loss: 0.0001092546372092329
step: 30, loss: 0.00036246070521883667
step: 40, loss: 5.502740532392636e-05
step: 50, loss: 9.27813962334767e-05
step: 60, loss: 9.790535841602832e-05
step: 70, loss: 8.851083839545026e-05
step: 80, loss: 0.011770102195441723
step: 90, loss: 3.818269760813564e-05
step: 100, loss: 0.00026690951199270785
step: 110, loss: 4.6757966629229486e-05
step: 120, loss: 9.411215432919562e-05
step: 130, loss: 0.00020406267140060663
step: 140, loss: 0.0002883262059185654
step: 150, loss: 0.00020095346553716809
step: 160, loss: 9.985377255361527e-05
step: 170, loss: 3.986648516729474e-05
step: 180, loss: 5.802436135127209e-05
step: 190, loss: 0.0034246372524648905
step: 200, loss: 6.717650103382766e-05
step: 210, loss: 0.04082884266972542
step: 220, loss: 0.02484671026468277
step: 230, loss: 0.00456473883241415
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9741282339707535, f1=0.9642058165548099,
best_f1=0.9620535714285715
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True

```

```

opt_level          : 02
cast_model_type    : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights     : True
loss_scale         : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2473it [00:07, 352.92it/s]
load_f1 = 0.96875
real_f1 = 0.9688888888888889
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:10, 422.03it/s]

```

### 4.3.2 DBLP-GoogleScholar - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Dirty/DBLP-GoogleScholar \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm distilbert \
  --fp16 \
  --save_model \
  --dk general \
  --dk product \
  --summarize \
  --da entry_swap \
  --da attr_del

```



```

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-GoogleScholar/.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Dirty/DBLP-GoogleScholar \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm distilbert \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
 /usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0  
 /usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`.

In PyTorch 1.1.0 and later, you should call them in the opposite order:

`optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

```

    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
    UserWarning)
step: 0, loss: 0.7936963438987732
step: 10, loss: 0.41962966322898865
step: 20, loss: 0.49493592977523804
step: 30, loss: 0.41830167174339294
step: 40, loss: 0.3851361870765686
step: 50, loss: 0.2393309324979782
step: 60, loss: 0.27312546968460083
step: 70, loss: 0.1996559202671051
step: 80, loss: 0.09821578115224838
step: 90, loss: 0.19627167284488678
step: 100, loss: 0.14204029738903046
step: 110, loss: 0.1552734673023224
step: 120, loss: 0.14445865154266357
step: 130, loss: 0.063322514295578
step: 140, loss: 0.09619558602571487
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 150, loss: 0.03337914124131203
step: 160, loss: 0.1968250423669815
step: 170, loss: 0.15449954569339752
step: 180, loss: 0.1071610152721405
step: 190, loss: 0.07170599699020386
step: 200, loss: 0.08880222588777542
step: 210, loss: 0.11479409039020538
step: 220, loss: 0.13485978543758392
step: 230, loss: 0.16851723194122314
step: 240, loss: 0.14749455451965332
step: 250, loss: 0.05607395991683006
step: 260, loss: 0.02883019670844078
step: 270, loss: 0.009182024747133255
step: 280, loss: 0.19289305806159973
step: 290, loss: 0.17398212850093842
step: 300, loss: 0.17527127265930176
step: 310, loss: 0.08523384481668472
step: 320, loss: 0.1277497559785843
step: 330, loss: 0.1367991417646408
step: 340, loss: 0.37245723605155945
step: 350, loss: 0.09675241261720657
step: 360, loss: 0.11194336414337158
step: 370, loss: 0.21989750862121582
step: 380, loss: 0.20828823745250702
step: 390, loss: 0.08088371902704239
step: 400, loss: 0.058991968631744385
step: 410, loss: 0.02540285885334015
step: 420, loss: 0.021081898361444473
step: 430, loss: 0.02884315885603428
step: 440, loss: 0.15472707152366638

```

```

step: 450, loss: 0.037710126489400864
step: 460, loss: 0.05804656445980072
step: 470, loss: 0.2504504323005676
step: 480, loss: 0.22784538567066193
step: 490, loss: 0.04863014444708824
step: 500, loss: 0.012534480541944504
step: 510, loss: 0.05827302485704422
step: 520, loss: 0.03389345109462738
step: 530, loss: 0.057232700288295746
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9173630454967502, f1=0.9105166051660517,
best_f1=0.9105166051660517
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11494620889425278
step: 10, loss: 0.1403370499610901
step: 20, loss: 0.19953250885009766
step: 30, loss: 0.0594879612326622
step: 40, loss: 0.005703662522137165
step: 50, loss: 0.04018014296889305
step: 60, loss: 0.1398332417011261
step: 70, loss: 0.1733178198337555
step: 80, loss: 0.014406402595341206
step: 90, loss: 0.011533195152878761
step: 100, loss: 0.3510872423648834
step: 110, loss: 0.036666955798864365
step: 120, loss: 0.09968393296003342
step: 130, loss: 0.01710224151611328
step: 140, loss: 0.010247284546494484
step: 150, loss: 0.0507306270301342
step: 160, loss: 0.044115468859672546
step: 170, loss: 0.06363724172115326
step: 180, loss: 0.0477924570441246
step: 190, loss: 0.12996365129947662
step: 200, loss: 0.1620020568370819
step: 210, loss: 0.02145320363342762
step: 220, loss: 0.1866859495639801
step: 230, loss: 0.029362112283706665
step: 240, loss: 0.20274867117404938
step: 250, loss: 0.035439860075712204
step: 260, loss: 0.0686890184879303

```

```

step: 270, loss: 0.06228838115930557
step: 280, loss: 0.13226786255836487
step: 290, loss: 0.06625385582447052
step: 300, loss: 0.14063431322574615
step: 310, loss: 0.08836189657449722
step: 320, loss: 0.050904568284749985
step: 330, loss: 0.05186525359749794
step: 340, loss: 0.003919404000043869
step: 350, loss: 0.0792841985821724
step: 360, loss: 0.1435719132423401
step: 370, loss: 0.018302071839571
step: 380, loss: 0.1226988136768341
step: 390, loss: 0.007794097065925598
step: 400, loss: 0.14078667759895325
step: 410, loss: 0.0009493673569522798
step: 420, loss: 0.05222293362021446
step: 430, loss: 0.04524340108036995
step: 440, loss: 0.010625340975821018
step: 450, loss: 0.042597446590662
step: 460, loss: 0.18467240035533905
step: 470, loss: 0.09588926285505295
step: 480, loss: 0.24884088337421417
step: 490, loss: 0.03386881947517395
step: 500, loss: 0.021895552054047585
step: 510, loss: 0.10169234871864319
step: 520, loss: 0.05245963856577873
step: 530, loss: 0.11888523399829865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9257722452743199, f1=0.9265588914549654,
best_f1=0.9265588914549654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02908424101769924
step: 10, loss: 0.08535312116146088
step: 20, loss: 0.1749960035085678
step: 30, loss: 0.18966266512870789
step: 40, loss: 0.03795095160603523
step: 50, loss: 0.04156513884663582
step: 60, loss: 0.007402051240205765
step: 70, loss: 0.015585925430059433
step: 80, loss: 0.011979312635958195

```

```
step: 90, loss: 0.1035805195569992
step: 100, loss: 0.011044860817492008
step: 110, loss: 0.011934695765376091
step: 120, loss: 0.11983411014080048
step: 130, loss: 0.027744043618440628
step: 140, loss: 0.02344980463385582
step: 150, loss: 0.03548314794898033
step: 160, loss: 0.02011403627693653
step: 170, loss: 0.009649290703237057
step: 180, loss: 0.07520808279514313
step: 190, loss: 0.016617335379123688
step: 200, loss: 0.021528039127588272
step: 210, loss: 0.06155693903565407
step: 220, loss: 0.04263759404420853
step: 230, loss: 0.04003728926181793
step: 240, loss: 0.011295394971966743
step: 250, loss: 0.021252920851111412
step: 260, loss: 0.006692197639495134
step: 270, loss: 0.007045181468129158
step: 280, loss: 0.02963402308523655
step: 290, loss: 0.02145116589963436
step: 300, loss: 0.08561072498559952
step: 310, loss: 0.10258718580007553
step: 320, loss: 0.14203883707523346
step: 330, loss: 0.016209768131375313
step: 340, loss: 0.0069502065889537334
step: 350, loss: 0.02434873767197132
step: 360, loss: 0.009845379739999771
step: 370, loss: 0.006214497610926628
step: 380, loss: 0.11474139243364334
step: 390, loss: 0.11423224210739136
step: 400, loss: 0.04461715370416641
step: 410, loss: 0.026551280170679092
step: 420, loss: 0.03758378326892853
step: 430, loss: 0.04097326472401619
step: 440, loss: 0.07305365055799484
step: 450, loss: 0.09969092905521393
step: 460, loss: 0.08410833030939102
step: 470, loss: 0.018375754356384277
step: 480, loss: 0.008021049201488495
step: 490, loss: 0.05226752534508705
step: 500, loss: 0.09811563789844513
step: 510, loss: 0.009054267778992653
step: 520, loss: 0.0039944881573319435
step: 530, loss: 0.14756228029727936
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
epoch 3: dev_f1=0.9244320815948077, f1=0.9274826789838336,  
best_f1=0.9265588914549654  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.006000754423439503  
step: 10, loss: 0.014461658895015717  
step: 20, loss: 0.0892593264579773  
step: 30, loss: 0.0013881037011742592  
step: 40, loss: 0.010594640858471394  
step: 50, loss: 0.060139868408441544  
step: 60, loss: 0.0514131523668766  
step: 70, loss: 0.0006750921602360904  
step: 80, loss: 0.00974820926785469  
step: 90, loss: 0.03479974716901779  
step: 100, loss: 0.07423538714647293  
step: 110, loss: 0.015138134360313416  
step: 120, loss: 0.0057422141544520855  
step: 130, loss: 0.07690643519163132  
step: 140, loss: 0.004302058834582567  
step: 150, loss: 0.0014076783554628491  
step: 160, loss: 0.015597863122820854  
step: 170, loss: 0.019479872658848763  
step: 180, loss: 0.007772103417664766  
step: 190, loss: 0.006760294083505869  
step: 200, loss: 0.001779177226126194  
step: 210, loss: 0.06163492798805237  
step: 220, loss: 0.010410592891275883  
step: 230, loss: 0.18109099566936493  
step: 240, loss: 0.006367028225213289  
step: 250, loss: 0.021458249539136887  
step: 260, loss: 0.1263851672410965  
step: 270, loss: 0.1112774983048439  
step: 280, loss: 0.0016009232494980097  
step: 290, loss: 0.02072257176041603  
step: 300, loss: 0.005561788100749254  
step: 310, loss: 0.0073399776592850685  
step: 320, loss: 0.1093408539891243  
step: 330, loss: 0.120623379945755  
step: 340, loss: 0.0288870669901371  
step: 350, loss: 0.0015392252244055271  
step: 360, loss: 0.06805402040481567  
step: 370, loss: 0.014528579078614712  
step: 380, loss: 0.0037635669577866793  
step: 390, loss: 0.02939577028155327  
step: 400, loss: 0.03174426034092903  
step: 410, loss: 0.0865633562207222  
step: 420, loss: 0.03275144100189209
```

```
step: 430, loss: 0.027568217366933823
step: 440, loss: 0.10232464224100113
step: 450, loss: 0.0044394079595804214
step: 460, loss: 0.0004674356314353645
step: 470, loss: 0.006260672118514776
step: 480, loss: 0.010436476208269596
step: 490, loss: 0.014758286997675896
step: 500, loss: 0.027177514508366585
step: 510, loss: 0.0673738643527031
step: 520, loss: 0.018877016380429268
step: 530, loss: 0.0008229490485973656
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.9225058004640372, f1=0.9244320815948077,
best_f1=0.9265588914549654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0070123556070029736
step: 10, loss: 0.014480345882475376
step: 20, loss: 0.023451190441846848
step: 30, loss: 0.06359418481588364
step: 40, loss: 0.02840297482907772
step: 50, loss: 0.012823662720620632
step: 60, loss: 0.0013809113297611475
step: 70, loss: 0.001032409374602139
step: 80, loss: 0.023885730654001236
step: 90, loss: 0.05147755518555641
step: 100, loss: 0.01673201099038124
step: 110, loss: 0.007336051668971777
step: 120, loss: 0.11932272464036942
step: 130, loss: 0.054140008985996246
step: 140, loss: 0.0028916841838508844
step: 150, loss: 0.008126871660351753
step: 160, loss: 0.017861323431134224
step: 170, loss: 0.07895070314407349
step: 180, loss: 0.007399868685752153
step: 190, loss: 0.00113185856025666
step: 200, loss: 0.0582817941904068
step: 210, loss: 0.04150447994470596
step: 220, loss: 0.0014864958357065916
step: 230, loss: 0.007904564030468464
step: 240, loss: 0.024375230073928833
step: 250, loss: 0.001138414372690022
step: 260, loss: 0.021209128201007843
step: 270, loss: 0.03924548253417015
step: 280, loss: 0.05790640041232109
```

```

step: 290, loss: 0.007948646321892738
step: 300, loss: 0.137750044465065
step: 310, loss: 0.007945573888719082
step: 320, loss: 0.008238410577178001
step: 330, loss: 0.0006676196935586631
step: 340, loss: 0.0006611694698221982
step: 350, loss: 0.06701786071062088
step: 360, loss: 0.012128079310059547
step: 370, loss: 0.00920180045068264
step: 380, loss: 0.020106077194213867
step: 390, loss: 0.01652473583817482
step: 400, loss: 0.002912259427830577
step: 410, loss: 0.0016975346952676773
step: 420, loss: 0.019869931042194366
step: 430, loss: 0.018970351666212082
step: 440, loss: 0.07758255302906036
step: 450, loss: 0.016687162220478058
step: 460, loss: 0.001801276346668601
step: 470, loss: 0.003205864690244198
step: 480, loss: 0.019562888890504837
step: 490, loss: 0.027787815779447556
step: 500, loss: 0.002581625245511532
step: 510, loss: 0.17816099524497986
step: 520, loss: 0.003809589659795165
step: 530, loss: 0.00885467417538166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.9244935543278084, f1=0.9279778393351801,
best_f1=0.9265588914549654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007960811257362366
step: 10, loss: 0.0017182754818350077
step: 20, loss: 0.02430451288819313
step: 30, loss: 0.03229523450136185
step: 40, loss: 0.014508897438645363
step: 50, loss: 0.0027784749399870634
step: 60, loss: 0.0071895238943398
step: 70, loss: 0.00970552209764719
step: 80, loss: 0.02610144205391407
step: 90, loss: 0.004348676186054945
step: 100, loss: 0.0009400340495631099
step: 110, loss: 0.026397647336125374
step: 120, loss: 0.0031482730992138386
step: 130, loss: 0.00502022635191679
step: 140, loss: 0.1747106909751892

```



```

step: 150, loss: 0.007200112100690603
step: 160, loss: 0.0009346937877126038
step: 170, loss: 0.0012524108169600368
step: 180, loss: 0.0027566049247980118
step: 190, loss: 0.0957195907831192
step: 200, loss: 0.0032677871640771627
step: 210, loss: 0.006619793362915516
step: 220, loss: 0.0007669540354982018
step: 230, loss: 0.007976330816745758
step: 240, loss: 0.002198402537032962
step: 250, loss: 0.00909676868468523
step: 260, loss: 0.0011447615688666701
step: 270, loss: 0.0005918850656598806
step: 280, loss: 0.008733512833714485
step: 290, loss: 0.0009471009252592921
step: 300, loss: 0.0016622249968349934
step: 310, loss: 0.008980220183730125
step: 320, loss: 0.07097068428993225
step: 330, loss: 0.0012684152461588383
step: 340, loss: 0.001299652038142085
step: 350, loss: 0.13548249006271362
step: 360, loss: 0.0029720321763306856
step: 370, loss: 0.14805452525615692
step: 380, loss: 0.10594606399536133
step: 390, loss: 0.0064527844078838825
step: 400, loss: 0.003654762636870146
step: 410, loss: 0.0017830987926572561
step: 420, loss: 0.09509523212909698
step: 430, loss: 0.006631460506469011
step: 440, loss: 0.11903629451990128
step: 450, loss: 0.001758099882863462
step: 460, loss: 0.0016871450934559107
step: 470, loss: 0.00110851158387959
step: 480, loss: 0.002535575535148382
step: 490, loss: 0.0014584959717467427
step: 500, loss: 0.0012261428637430072
step: 510, loss: 0.00015605833323206753
step: 520, loss: 0.014581151306629181
step: 530, loss: 0.0017088266322389245
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9185667752442996, f1=0.9205759405480723,
best_f1=0.9265588914549654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0054154605604708195

```

step: 10, loss: 0.010943492874503136  
step: 20, loss: 0.016804108396172523  
step: 30, loss: 0.005804252810776234  
step: 40, loss: 0.0007642176933586597  
step: 50, loss: 0.0007109592552296817  
step: 60, loss: 0.0015721169766038656  
step: 70, loss: 0.004837254993617535  
step: 80, loss: 0.0004734335234388709  
step: 90, loss: 0.0009245309047400951  
step: 100, loss: 0.002761934185400605  
step: 110, loss: 0.004693288821727037  
step: 120, loss: 0.08909320831298828  
step: 130, loss: 0.03642377257347107  
step: 140, loss: 0.0011396821355447173  
step: 150, loss: 0.0021020814310759306  
step: 160, loss: 0.0006169883999973536  
step: 170, loss: 0.001094173640012741  
step: 180, loss: 0.08378219604492188  
step: 190, loss: 0.0007097076741047204  
step: 200, loss: 0.0015314951306208968  
step: 210, loss: 0.0008615657570771873  
step: 220, loss: 0.02529910020530224  
step: 230, loss: 0.0005341747892089188  
step: 240, loss: 0.004826856777071953  
step: 250, loss: 0.0030874768272042274  
step: 260, loss: 0.002146882237866521  
step: 270, loss: 0.014876540750265121  
step: 280, loss: 0.0017297635786235332  
step: 290, loss: 0.023943325504660606  
step: 300, loss: 0.002409170148894191  
step: 310, loss: 0.0128647331148386  
step: 320, loss: 0.10488973557949066  
step: 330, loss: 0.00470641627907753  
step: 340, loss: 0.010191680863499641  
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0  
step: 350, loss: 0.00024890160420909524  
step: 360, loss: 0.004713854286819696  
step: 370, loss: 0.011987131088972092  
step: 380, loss: 0.009068277664482594  
step: 390, loss: 0.0005782697699032724  
step: 400, loss: 7.30328683857806e-05  
step: 410, loss: 0.002332883421331644  
step: 420, loss: 0.0001612109481357038  
step: 430, loss: 4.860903572989628e-05  
step: 440, loss: 0.00020349207625258714  
step: 450, loss: 0.06832513213157654  
step: 460, loss: 0.002557442057877779  
step: 470, loss: 0.0020156896207481623

```

step: 480, loss: 0.007024228107184172
step: 490, loss: 0.0321807824075222
step: 500, loss: 0.00027926830807700753
step: 510, loss: 0.007868695072829723
step: 520, loss: 0.010158873163163662
step: 530, loss: 6.070718518458307e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9223573433115062, f1=0.9245901639344263,
best_f1=0.9265588914549654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010897611500695348
step: 10, loss: 0.0026385702658444643
step: 20, loss: 0.020080674439668655
step: 30, loss: 0.019596362486481667
step: 40, loss: 0.00039533336530439556
step: 50, loss: 0.0002633127151057124
step: 60, loss: 0.006115685682743788
step: 70, loss: 0.03271960839629173
step: 80, loss: 5.493028584169224e-05
step: 90, loss: 0.0002287227107444778
step: 100, loss: 0.005514612887054682
step: 110, loss: 0.0006039325962774456
step: 120, loss: 0.0157322958111763
step: 130, loss: 0.0019470229744911194
step: 140, loss: 4.854223152506165e-05
step: 150, loss: 0.004671567119657993
step: 160, loss: 0.0011368617415428162
step: 170, loss: 0.019228016957640648
step: 180, loss: 0.0014645365299656987
step: 190, loss: 0.01670987717807293
step: 200, loss: 0.001947434269823134
step: 210, loss: 0.05379006266593933
step: 220, loss: 0.003836193820461631
step: 230, loss: 0.0385357104241848
step: 240, loss: 0.00035672116791829467
step: 250, loss: 0.008540123701095581
step: 260, loss: 0.15753577649593353
step: 270, loss: 0.00023675835109315813
step: 280, loss: 0.005186921451240778
step: 290, loss: 0.043865516781806946
step: 300, loss: 0.0008096657111309469
step: 310, loss: 0.09179271012544632
step: 320, loss: 0.012486578896641731
step: 330, loss: 0.07281414419412613

```

```

step: 340, loss: 0.002029229188337922
step: 350, loss: 0.011477235704660416
step: 360, loss: 0.011602935381233692
step: 370, loss: 0.006229951046407223
step: 380, loss: 0.0016524603124707937
step: 390, loss: 0.0019036623416468501
step: 400, loss: 0.06318160891532898
step: 410, loss: 0.001936628483235836
step: 420, loss: 0.0014346307143568993
step: 430, loss: 0.0017883613472804427
step: 440, loss: 0.02022385410964489
step: 450, loss: 0.004525875207036734
step: 460, loss: 0.004369200672954321
step: 470, loss: 0.0006737548392266035
step: 480, loss: 0.004231403581798077
step: 490, loss: 0.00047367517254315317
step: 500, loss: 0.0029226024635136127
step: 510, loss: 0.0006791446940042078
step: 520, loss: 0.0008826153934933245
step: 530, loss: 0.00011305641237413511
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9222222222222223, f1=0.9237875288683602,
best_f1=0.9265588914549654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016324471216648817
step: 10, loss: 0.00031970409327186644
step: 20, loss: 0.00011339491902617738
step: 30, loss: 0.005393680185079575
step: 40, loss: 0.01609095185995102
step: 50, loss: 0.0003864118771161884
step: 60, loss: 0.0002172249078284949
step: 70, loss: 0.12919799983501434
step: 80, loss: 0.0029763104394078255
step: 90, loss: 0.040009208023548126
step: 100, loss: 0.0010424890788272023
step: 110, loss: 0.001364495838060975
step: 120, loss: 0.0030120075680315495
step: 130, loss: 0.00023414974566549063
step: 140, loss: 0.04508921504020691
step: 150, loss: 0.00018664127856027335
step: 160, loss: 0.00011557209654711187
step: 170, loss: 9.819628030527383e-05
step: 180, loss: 0.00028346184990368783
step: 190, loss: 0.015712235122919083

```

```
step: 200, loss: 0.00458791246637702
step: 210, loss: 0.00014940000255592167
step: 220, loss: 0.0013652635971084237
step: 230, loss: 0.00022195052588358521
step: 240, loss: 0.0161714106798172
step: 250, loss: 0.016329504549503326
step: 260, loss: 0.0009035113034769893
step: 270, loss: 0.06466931849718094
step: 280, loss: 0.004474470857530832
step: 290, loss: 0.00014242609904613346
step: 300, loss: 0.000107243460661266
step: 310, loss: 0.00017023205873556435
step: 320, loss: 0.0002387556160101667
step: 330, loss: 0.016935668885707855
step: 340, loss: 0.004702393896877766
step: 350, loss: 0.00037878763396292925
step: 360, loss: 0.010178917087614536
step: 370, loss: 0.0006659788195975125
step: 380, loss: 0.014423416927456856
step: 390, loss: 0.001378739601932466
step: 400, loss: 0.0007929243729449809
step: 410, loss: 0.00014417928468901664
step: 420, loss: 0.001725618029013276
step: 430, loss: 0.0001574769412400201
step: 440, loss: 8.568174962420017e-05
step: 450, loss: 0.00024756425409577787
step: 460, loss: 0.0049840183928608894
step: 470, loss: 0.00023232768580783159
step: 480, loss: 0.0005078427493572235
step: 490, loss: 0.021151103079319
step: 500, loss: 0.007688011974096298
step: 510, loss: 0.0006756203947588801
step: 520, loss: 0.00012510694796219468
step: 530, loss: 0.07890338450670242
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9206945096198967, f1=0.9214953271028037,
best_f1=0.9265588914549654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13231432437896729
step: 10, loss: 0.00017485898570157588
step: 20, loss: 0.0025254162028431892
step: 30, loss: 0.00018323813856113702
step: 40, loss: 0.000801789341494441
step: 50, loss: 0.0003188005939591676
```

step: 60, loss: 0.00031445990316569805  
step: 70, loss: 0.0035202407743781805  
step: 80, loss: 0.0027847038581967354  
step: 90, loss: 0.012063469737768173  
step: 100, loss: 0.0010497656185179949  
step: 110, loss: 0.0011952570639550686  
step: 120, loss: 0.017385773360729218  
step: 130, loss: 5.839763252879493e-05  
step: 140, loss: 0.00036742104566656053  
step: 150, loss: 0.00029270153027027845  
step: 160, loss: 0.00021275304607115686  
step: 170, loss: 0.00012143004278186709  
step: 180, loss: 0.002744101220741868  
step: 190, loss: 0.0005300953052937984  
step: 200, loss: 0.00014752506103832275  
step: 210, loss: 8.193671965273097e-05  
step: 220, loss: 0.0001840934419305995  
step: 230, loss: 0.00022916414309293032  
step: 240, loss: 4.326382986619137e-05  
step: 250, loss: 0.00023781503841746598  
step: 260, loss: 0.0043657319620251656  
step: 270, loss: 5.94233424635604e-05  
step: 280, loss: 6.844644667580724e-05  
step: 290, loss: 0.00016640766989439726  
step: 300, loss: 0.0002784670505207032  
step: 310, loss: 0.0006494391127489507  
step: 320, loss: 0.0009503628825768828  
step: 330, loss: 0.00840876903384924  
step: 340, loss: 0.00018030389037448913  
step: 350, loss: 0.0007969783619046211  
step: 360, loss: 0.038934603333473206  
step: 370, loss: 0.0016041623894125223  
step: 380, loss: 0.010222677141427994  
step: 390, loss: 0.00015973545669112355  
step: 400, loss: 0.00020548062457237393  
step: 410, loss: 0.00026590144261717796  
step: 420, loss: 0.0017164561431854963  
step: 430, loss: 0.004269061144441366  
step: 440, loss: 0.00010865938384085894  
step: 450, loss: 0.03265155479311943  
step: 460, loss: 0.0028129371348768473  
step: 470, loss: 0.0001637004315853119  
step: 480, loss: 0.038798749446868896  
step: 490, loss: 0.01824372634291649  
step: 500, loss: 0.028563067317008972  
step: 510, loss: 0.00024617029703222215  
step: 520, loss: 0.0004920004867017269  
step: 530, loss: 0.00126290088519454

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9219261337073399, f1=0.9157303370786517,
best_f1=0.9265588914549654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005340244970284402
step: 10, loss: 0.006196728907525539
step: 20, loss: 0.00024821836268529296
step: 30, loss: 0.0006248197751119733
step: 40, loss: 0.00035967736039310694
step: 50, loss: 0.0022974146995693445
step: 60, loss: 0.0004917486221529543
step: 70, loss: 0.00016200856771320105
step: 80, loss: 0.0001906051329569891
step: 90, loss: 0.0002746685058809817
step: 100, loss: 0.01773850992321968
step: 110, loss: 0.00011440224625403062
step: 120, loss: 0.0004447940445970744
step: 130, loss: 0.0001586291764397174
step: 140, loss: 8.517177047906443e-05
step: 150, loss: 0.001216176780872047
step: 160, loss: 0.0003063042531721294
step: 170, loss: 0.0016309291822835803
step: 180, loss: 0.0006815191591158509
step: 190, loss: 0.0005796017940156162
step: 200, loss: 0.004891750402748585
step: 210, loss: 0.000909570837393403
step: 220, loss: 0.004981475882232189
step: 230, loss: 0.0011412808671593666
step: 240, loss: 0.0031756602693349123
step: 250, loss: 0.0009073962573893368
step: 260, loss: 6.008456330164336e-05
step: 270, loss: 0.00391797861084342
step: 280, loss: 0.001917596673592925
step: 290, loss: 0.0005817866185680032
step: 300, loss: 0.12414336949586868
step: 310, loss: 0.0004228315665386617
step: 320, loss: 0.002064438071101904
step: 330, loss: 0.0015918652061372995
step: 340, loss: 0.0005027435254305601
step: 350, loss: 0.0005448046722449362
step: 360, loss: 0.0013628995511680841
step: 370, loss: 0.00014257944712881
step: 380, loss: 4.417201489559375e-05
step: 390, loss: 0.004529525060206652
```

```

step: 400, loss: 6.675469921901822e-05
step: 410, loss: 0.005721334833651781
step: 420, loss: 0.00027898672851733863
step: 430, loss: 7.492449367418885e-05
step: 440, loss: 6.137161108199507e-05
step: 450, loss: 0.0006529712118208408
step: 460, loss: 0.023755956441164017
step: 470, loss: 0.0037391234654933214
step: 480, loss: 0.00035020490759052336
step: 490, loss: 0.0028629512526094913
step: 500, loss: 0.00013697089161723852
step: 510, loss: 2.694430440897122e-05
step: 520, loss: 8.80492152646184e-05
step: 530, loss: 0.00013987223792355508
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9198113207547169, f1=0.9187145557655955,
best_f1=0.9265588914549654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00045992308878339827
step: 10, loss: 6.519832095364109e-05
step: 20, loss: 4.6203844249248505e-05
step: 30, loss: 0.06547340005636215
step: 40, loss: 0.0001700037973932922
step: 50, loss: 0.00030437568784691393
step: 60, loss: 0.0003171337302774191
step: 70, loss: 0.0010373394470661879
step: 80, loss: 0.001686933683231473
step: 90, loss: 0.0001606261939741671
step: 100, loss: 0.0011101665440946817
step: 110, loss: 0.08822201937437057
step: 120, loss: 0.0020077659282833338
step: 130, loss: 0.032228220254182816
step: 140, loss: 0.00013577153731603175
step: 150, loss: 0.005032742861658335
step: 160, loss: 0.000987717416137457
step: 170, loss: 0.002268875250592828
step: 180, loss: 0.0013955067843198776
step: 190, loss: 0.002029197756201029
step: 200, loss: 0.0005159459542483091
step: 210, loss: 0.0004962276434525847
step: 220, loss: 0.0005390022997744381
step: 230, loss: 0.0056546274572610855
step: 240, loss: 0.0011342742945998907
step: 250, loss: 0.0003051870735362172

```



```

step: 260, loss: 0.002314299112185836
step: 270, loss: 0.0073648844845592976
step: 280, loss: 0.00016386076458729804
step: 290, loss: 0.00513302581384778
step: 300, loss: 0.16429823637008667
step: 310, loss: 0.0013258001999929547
step: 320, loss: 0.00011053845810238272
step: 330, loss: 0.0017314208671450615
step: 340, loss: 0.008533462882041931
step: 350, loss: 0.0031974329613149166
step: 360, loss: 0.0013246700400486588
step: 370, loss: 0.00033924353192560375
step: 380, loss: 0.00016168670845218003
step: 390, loss: 0.0003257899370510131
step: 400, loss: 0.00017235793347936124
step: 410, loss: 0.0006693213945254683
step: 420, loss: 0.0014352357247844338
step: 430, loss: 0.0008273666608147323
step: 440, loss: 0.0017475895583629608
step: 450, loss: 0.0008387608686462045
step: 460, loss: 0.0011955631198361516
step: 470, loss: 0.14575956761837006
step: 480, loss: 0.00021929075592197478
step: 490, loss: 0.00021940410078968853
step: 500, loss: 0.009911310859024525
step: 510, loss: 0.0004085579130332917
step: 520, loss: 0.006584870629012585
step: 530, loss: 0.0007271904032677412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9222170470423847, f1=0.9251637043966324,
best_f1=0.9265588914549654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00694855023175478
step: 10, loss: 0.0001825483050197363
step: 20, loss: 0.048297375440597534
step: 30, loss: 0.00032951199682429433
step: 40, loss: 0.0019660957623273134
step: 50, loss: 0.0010062003275379539
step: 60, loss: 0.00018016243120655417
step: 70, loss: 9.100315219257027e-05
step: 80, loss: 0.0002965500170830637
step: 90, loss: 0.004862107802182436
step: 100, loss: 2.7722475351765752e-05
step: 110, loss: 0.00492493761703372

```

```
step: 120, loss: 0.0003614417219068855
step: 130, loss: 0.0022400259040296078
step: 140, loss: 5.31430596311111e-05
step: 150, loss: 0.0002827310818247497
step: 160, loss: 7.617802475579083e-05
step: 170, loss: 0.00022964955132920295
step: 180, loss: 6.21294166194275e-05
step: 190, loss: 0.0003861206932924688
step: 200, loss: 0.0013515338068827987
step: 210, loss: 0.0022234844509512186
step: 220, loss: 0.0001401925546815619
step: 230, loss: 0.00774258142337203
step: 240, loss: 0.00014033936895430088
step: 250, loss: 0.0877535417675972
step: 260, loss: 0.0018067299388349056
step: 270, loss: 0.0023463491816073656
step: 280, loss: 0.0030197915621101856
step: 290, loss: 0.00020976261293981224
step: 300, loss: 0.006923042703419924
step: 310, loss: 0.0003759117389563471
step: 320, loss: 5.121835783938877e-05
step: 330, loss: 0.0008599022985436022
step: 340, loss: 0.00600869907066226
step: 350, loss: 0.011726508848369122
step: 360, loss: 7.258017285494134e-05
step: 370, loss: 0.0004609478055499494
step: 380, loss: 0.0017482335679233074
step: 390, loss: 0.00030252133728936315
step: 400, loss: 8.340272324858233e-05
step: 410, loss: 7.214875222416595e-05
step: 420, loss: 0.0002622803149279207
step: 430, loss: 0.0006569543038494885
step: 440, loss: 0.0005473954370245337
step: 450, loss: 0.0004094892356079072
step: 460, loss: 2.4131966711138375e-05
step: 470, loss: 0.0003430589276831597
step: 480, loss: 0.00177648919634521
step: 490, loss: 0.001182169420644641
step: 500, loss: 0.002538459375500679
step: 510, loss: 5.2442461310420185e-05
step: 520, loss: 0.0004518919449765235
step: 530, loss: 0.00112536468077451
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.918918918918919, f1=0.920205319645357,
best_f1=0.9265588914549654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.00013954422320239246
step: 10, loss: 0.0007408473175019026
step: 20, loss: 0.00016569302533753216
step: 30, loss: 0.0005924658034928143
step: 40, loss: 0.0024491488002240658
step: 50, loss: 0.001325408462435007
step: 60, loss: 2.690709152375348e-05
step: 70, loss: 8.801752119325101e-05
step: 80, loss: 9.867355402093381e-05
step: 90, loss: 0.0004710771609097719
step: 100, loss: 0.00037973394501022995
step: 110, loss: 3.101956099271774e-05
step: 120, loss: 0.0017873819451779127
step: 130, loss: 0.0029141732957214117
step: 140, loss: 0.025048334151506424
step: 150, loss: 0.0018507768400013447
step: 160, loss: 0.0007383242482319474
step: 170, loss: 0.006005249917507172
step: 180, loss: 0.00015501018788199872
step: 190, loss: 0.0003456526901572943
step: 200, loss: 0.00012830087507609278
step: 210, loss: 0.00036661443300545216
step: 220, loss: 5.9936228353763e-05
step: 230, loss: 0.0026541573461145163
step: 240, loss: 0.014939848333597183
step: 250, loss: 0.0003298194205854088
step: 260, loss: 0.00021107376960571855
step: 270, loss: 0.001176981721073389
step: 280, loss: 0.00029096537036821246
step: 290, loss: 0.00012846445315517485
step: 300, loss: 0.001341387745924294
step: 310, loss: 2.8642767574638128e-05
step: 320, loss: 0.001596097950823605
step: 330, loss: 0.0005571295041590929
step: 340, loss: 0.002779265632852912
step: 350, loss: 0.002441229997202754
step: 360, loss: 0.014583843760192394
step: 370, loss: 0.025436008349061012
step: 380, loss: 0.11840558797121048
step: 390, loss: 0.0006363852298818529
step: 400, loss: 0.0004274248785804957
step: 410, loss: 0.00012094851263100281
step: 420, loss: 0.0005870875320397317
step: 430, loss: 6.12244475632906e-05
step: 440, loss: 0.00024719734210520983
step: 450, loss: 6.494518311228603e-05
```

```

step: 460, loss: 0.005819342564791441
step: 470, loss: 6.755764479748905e-05
step: 480, loss: 7.190743053797632e-05
step: 490, loss: 0.003887494094669819
step: 500, loss: 0.0006017790292389691
step: 510, loss: 8.042420085985214e-05
step: 520, loss: 4.6263910917332396e-05
step: 530, loss: 0.0006790503393858671
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.923005132991134, f1=0.9195509822263798,
best_f1=0.9265588914549654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.438075747108087e-05
step: 10, loss: 0.0001256104587810114
step: 20, loss: 0.0005438182852230966
step: 30, loss: 0.00024957029381766915
step: 40, loss: 0.00014071336772758514
step: 50, loss: 4.3075269786641e-05
step: 60, loss: 5.244654676062055e-05
step: 70, loss: 0.00026073798653669655
step: 80, loss: 0.0001422139466740191
step: 90, loss: 1.750494629959576e-05
step: 100, loss: 4.703173544839956e-05
step: 110, loss: 0.013114433735609055
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 120, loss: 0.0006545516080223024
step: 130, loss: 0.0009741291869431734
step: 140, loss: 3.975250365328975e-05
step: 150, loss: 0.011316128075122833
step: 160, loss: 8.07786564109847e-05
step: 170, loss: 0.0002878880768548697
step: 180, loss: 0.0005931832711212337
step: 190, loss: 0.00038477114867419004
step: 200, loss: 7.677920075366274e-05
step: 210, loss: 0.0015476180706173182
step: 220, loss: 8.90018927748315e-05
step: 230, loss: 0.0008327745599672198
step: 240, loss: 0.00032085354905575514
step: 250, loss: 0.0015820176340639591
step: 260, loss: 0.01088433712720871
step: 270, loss: 0.0009800158441066742
step: 280, loss: 0.0005955661181360483
step: 290, loss: 0.0006643860833719373
step: 300, loss: 0.0006597895408049226

```

```

step: 310, loss: 0.0006564278155565262
step: 320, loss: 0.003803553991019726
step: 330, loss: 0.00016887817764654756
step: 340, loss: 0.009842115454375744
step: 350, loss: 0.00023494228662457317
step: 360, loss: 0.0027828074526041746
step: 370, loss: 0.0006678006611764431
step: 380, loss: 0.0002963048464152962
step: 390, loss: 0.00013290959759615362
step: 400, loss: 0.0010142629034817219
step: 410, loss: 0.0003965610230807215
step: 420, loss: 0.0002977123949676752
step: 430, loss: 0.00010956113692373037
step: 440, loss: 0.009686457924544811
step: 450, loss: 0.006551739759743214
step: 460, loss: 0.0001588022569194436
step: 470, loss: 6.37205594102852e-05
step: 480, loss: 0.0007599206292070448
step: 490, loss: 0.0001920002105180174
step: 500, loss: 2.6749015887617134e-05
step: 510, loss: 3.1347306503448635e-05
step: 520, loss: 0.00010149212903343141
step: 530, loss: 7.269605703186244e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9218530650444547, f1=0.9226441631504922,
best_f1=0.9265588914549654
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True

```

```

loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:14, 403.25it/s]
load_f1 = 0.9253456221198156
real_f1 = 0.92201199815413
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:10, 426.29it/s]

```