EMedium_90_1_2_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/Snippext_public
     %cd Snippext_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 | 8.48 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 80.8 MB/s
Collecting numpy==1.19.2
  Downloading numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
     |                        | 14.5 MB 14.1 MB/s
Collecting regex==2019.12.20
  Downloading regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
     |                        | 689 kB 53.2 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 75.4 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 63.6 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 | 11.51 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 30.8 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 75.1 MB/s
Collecting torch==1.4.0
  Downloading torch-1.4.0-cp37-cp37m-manylinux1_x86_64.whl (753.4 MB)
     |                      | 753.4 MB 7.1 kB/s
Collecting tqdm==4.41.0
  Downloading tqdm-4.41.0-py2.py3-none-any.whl (56 kB)
     |                      | 56 kB 5.4 MB/s
Collecting transformers==3.1.0
  Downloading transformers-3.1.0-py3-none-any.whl (884 kB)
     |                      | 884 kB 62.0 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 58.7 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 48.0 MB/s
Collecting catalogue<1.1.0,>=0.0.7
  Downloading catalogue-1.0.0-py2.py3-none-any.whl (7.7 kB)
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)
Collecting srsly<1.1.0,>=0.1.0
  Downloading srsly-1.0.5-cp37-cp37m-manylinux2014_x86_64.whl (184 kB)
     |                    | 184 kB 65.2 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)
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: 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 blis<0.5.0,>=0.4.0
  Downloading blis-0.4.1-cp37-cp37m-manylinux1_x86_64.whl (3.7 MB)
     |                    | 3.7 MB 52.7 MB/s
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: 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)
Requirement already satisfied: packaging in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (21.3)
Collecting sacremoses
  Downloading sacremoses-0.0.53.tar.gz (880 kB)
```

```
         |                    | 880 kB 46.8 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 55.6 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: 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: 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: 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: 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: 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: 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: 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: 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=1449921
sha256=bf9839a3c36dae051124f2066c7fc492aac34545391309c0e126499b4f685d3b
  Stored in directory: /root/.cache/pip/wheels/48/8b/7f/473521e0c731c6566d631b28
1f323842bbda9bd819eb9a3ead
  Building wheel for sacremoses (setup.py) … done
  Created wheel for sacremoses: filename=sacremoses-0.0.53-py3-none-any.whl
size=895260
```

```
sha256=5135eab09e95bb444d80f14b12fe92720f4264ce70d672659bc676e8bf56aaeb
  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 | 9.73 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 writeable
Created temporary directory: /tmp/pip-ephem-wheel-cache-tv20qbrx
Created temporary directory: /tmp/pip-req-tracker-ow2nldfb
Initialized build tracking at /tmp/pip-req-tracker-ow2nldfb
Created build tracker: /tmp/pip-req-tracker-ow2nldfb
Entered build tracker: /tmp/pip-req-tracker-ow2nldfb
```

```
Created temporary directory: /tmp/pip-install-4s5s358g
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-b03xu3pu
  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-
ow2nldfb'
    Running setup.py (path:/tmp/pip-req-build-b03xu3pu/setup.py) egg_info for
package from file:///content/ditto/apex
    Created temporary directory: /tmp/pip-pip-egg-info-ljzvk6gw
    Running command python setup.py egg_info


    torch.__version__  = 1.4.0


    running egg_info
    creating /tmp/pip-pip-egg-info-ljzvk6gw/apex.egg-info
    writing /tmp/pip-pip-egg-info-ljzvk6gw/apex.egg-info/PKG-INFO
    writing dependency_links to /tmp/pip-pip-egg-info-ljzvk6gw/apex.egg-
info/dependency_links.txt
    writing top-level names to /tmp/pip-pip-egg-info-ljzvk6gw/apex.egg-
info/top_level.txt
    writing manifest file '/tmp/pip-pip-egg-info-ljzvk6gw/apex.egg-
info/SOURCES.txt'
    adding license file 'LICENSE'
    writing manifest file '/tmp/pip-pip-egg-info-ljzvk6gw/apex.egg-
info/SOURCES.txt'
  Source in /tmp/pip-req-build-b03xu3pu 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-ow2nldfb'
Created temporary directory: /tmp/pip-unpack-i14uk8w9
Building wheels for collected packages: apex
  Created temporary directory: /tmp/pip-wheel-633z8qgl
  Building wheel for apex (setup.py) …   Destination directory:
/tmp/pip-wheel-633z8qgl
  Running command /usr/bin/python3 -u -c 'import io, os, sys, setuptools,
tokenize; sys.argv[0] = '"'"'/tmp/pip-req-build-b03xu3pu/setup.py'"'"';
```

```
__file__='"'"'/tmp/pip-req-build-b03xu3pu/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-633z8qgl


  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/_batchsampler.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_pipelining_with_i
nterleaving.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_pipelining.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_pipelining_withou
t_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/_batchsampler.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_pipelin
ing_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_pipelining.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_pipelin
ing_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-633z8qgl/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/_batchsampler.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=11f70ce5fb1f9c859c7d3f4b7493293b38646ade8c22a7dde0be852509a7e3bb
  Stored in directory: /tmp/pip-ephem-wheel-cache-
tv20qbrx/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-ow2nldfb'
/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 28.3 MB/s
Collecting awscli
  Downloading awscli-1.25.46-py3-none-any.whl (3.9 MB)
     | | 3.9 MB 62.7 MB/s
Collecting s3transfer<0.7.0,>=0.6.0
  Downloading s3transfer-0.6.0-py3-none-any.whl (79 kB)
     | | 79 kB 9.8 MB/s
Requirement already satisfied: PyYAML<5.5,>=3.10 in
/usr/local/lib/python3.7/dist-packages (from awscli) (3.13)
Collecting rsa<4.8,>=3.1.2
  Downloading rsa-4.7.2-py3-none-any.whl (34 kB)
Collecting botocore==1.27.46
  Downloading botocore-1.27.46-py3-none-any.whl (9.0 MB)
     | | 9.0 MB 63.1 MB/s
Collecting colorama<0.4.5,>=0.2.5
  Downloading colorama-0.4.4-py2.py3-none-any.whl (16 kB)
Collecting docutils<0.17,>=0.10
  Downloading docutils-0.16-py2.py3-none-any.whl (548 kB)
     | | 548 kB 52.3 MB/s
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 | 17.31 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/EMedium_90_1_2/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, 378kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 348kB/s]
Downloading: 100% 268M/268M [00:05<00:00, 45.4MB/s]
step: 0, loss: 0.8683480024337769
epoch 1: dev_f1=0.3404255319148936, f1=0.22641509433962267,
best_f1=0.22641509433962267
step: 0, loss: 0.4214494824409485
epoch 2: dev_f1=0.4166666666666667, f1=0.3142857142857143,
best_f1=0.3142857142857143
step: 0, loss: 0.34317490458488464
epoch 3: dev_f1=0.4117647058823529, f1=0.32911392405063294,
best_f1=0.3142857142857143
```

```
step: 0, loss: 0.3646322786808014
epoch 4: dev_f1=0.46153846153846156, f1=0.3561643835616438,
best_f1=0.3561643835616438
step: 0, loss: 0.2650502026081085
epoch 5: dev_f1=0.4864864864864865, f1=0.3548387096774194,
best_f1=0.3548387096774194
step: 0, loss: 0.2725885808467865
epoch 6: dev_f1=0.5454545454545455, f1=0.41379310344827586,
best_f1=0.41379310344827586
step: 0, loss: 0.26114651560783386
epoch 7: dev_f1=0.5454545454545454, f1=0.44897959183673464,
best_f1=0.41379310344827586
step: 0, loss: 0.32936981320381165
epoch 8: dev_f1=0.5294117647058824, f1=0.43999999999999995,
best_f1=0.41379310344827586
step: 0, loss: 0.1432206779718399
epoch 9: dev_f1=0.5806451612903226, f1=0.48888888888888893,
best_f1=0.48888888888888893
step: 0, loss: 0.2184714823961258
epoch 10: dev_f1=0.5714285714285714, f1=0.5714285714285714,
best_f1=0.48888888888888893
step: 0, loss: 0.10338044911623001
epoch 11: dev_f1=0.5833333333333334, f1=0.5517241379310344,
best_f1=0.5517241379310344
step: 0, loss: 0.13214588165283203
epoch 12: dev_f1=0.6086956521739131, f1=0.5714285714285714,
best_f1=0.5714285714285714
step: 0, loss: 0.07029344886541367
epoch 13: dev_f1=0.6470588235294117, f1=0.4878048780487805,
best_f1=0.4878048780487805
step: 0, loss: 0.09955426305532455
epoch 14: dev_f1=0.6470588235294117, f1=0.4615384615384615,
best_f1=0.4878048780487805
step: 0, loss: 0.1417129635810852
epoch 15: dev_f1=0.6470588235294117, f1=0.4615384615384615,
best_f1=0.4878048780487805
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 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: 0.8176185488700867
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4759708642959595
step: 20, loss: 0.5678945183753967
step: 30, loss: 0.35955947637557983
step: 40, loss: 0.12421838939189911
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.16103008389472961
step: 60, loss: 0.10938681662082672
step: 70, loss: 0.05522783845663071
step: 80, loss: 0.17497046291828156
step: 90, loss: 0.10162188112735748
step: 100, loss: 0.031959887593984604
step: 110, loss: 0.16393554210662842
step: 120, loss: 0.015326429158449173
step: 130, loss: 0.014776282012462616
step: 140, loss: 0.07007251679897308
step: 150, loss: 0.0926583781838417
step: 160, loss: 0.065201535820961
step: 170, loss: 0.003934954293072224
step: 180, loss: 0.01152989361435175
step: 190, loss: 0.011793004348874092
step: 200, loss: 0.021241772919893265
step: 210, loss: 0.040823791176080704
step: 220, loss: 0.038236863911151886
step: 230, loss: 0.029152698814868927
```

```
/usr/local/lib/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=0.972972972972973, f1=0.971815107102593,
best_f1=0.971815107102593
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013992472551763058
step: 10, loss: 0.0007809872040525079
step: 20, loss: 0.00803239643573761
step: 30, loss: 0.012387997470796108
step: 40, loss: 0.006960437633097172
step: 50, loss: 0.010815294459462166
step: 60, loss: 0.004256937652826309
step: 70, loss: 0.019115779548883438
step: 80, loss: 0.004621107131242752
step: 90, loss: 0.0035964243579655886
step: 100, loss: 0.04073613882064819
step: 110, loss: 0.01813620515167713
step: 120, loss: 0.003991374745965004
step: 130, loss: 0.0067537869326770306
step: 140, loss: 0.02541647106409073
step: 150, loss: 0.011175209656357765
step: 160, loss: 0.004342670552432537
step: 170, loss: 0.04165960103273392
step: 180, loss: 0.0476326160132885
step: 190, loss: 0.15558922290802002
step: 200, loss: 0.011513395234942436
step: 210, loss: 0.10939384251832962
step: 220, loss: 0.008797864429652691
step: 230, loss: 0.07159380614757538
/usr/local/lib/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=0.9778270509977827, 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: 0.0873931422829628
step: 10, loss: 0.006127079948782921
step: 20, loss: 0.02532586082816124
step: 30, loss: 0.008479340933263302
step: 40, loss: 0.030930083245038986
step: 50, loss: 0.003620629431679845
step: 60, loss: 0.07992493361234665
step: 70, loss: 0.00092599744675681
step: 80, loss: 0.0012440808350220323
step: 90, loss: 0.0014398450730368495
step: 100, loss: 0.038288746029138565
step: 110, loss: 0.11988629400730133
step: 120, loss: 0.004794810898602009
step: 130, loss: 0.0014489712193608284
step: 140, loss: 0.0027294026222079992
step: 150, loss: 0.0019192079780623317
step: 160, loss: 0.004514266271144152
step: 170, loss: 0.005918827839195728
step: 180, loss: 0.0010687687899917364
step: 190, loss: 0.006124622188508511
step: 200, loss: 0.00080401252489537
step: 210, loss: 0.1634681224822998
step: 220, loss: 0.008991874754428864
step: 230, loss: 0.07595381885766983
/usr/local/lib/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=0.9864559819413092, f1=0.9785310734463276,
best_f1=0.9785310734463276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005347329657524824
step: 10, loss: 0.0009201959474012256
step: 20, loss: 0.0011265033390372992
step: 30, loss: 0.0006914365221746266
step: 40, loss: 0.08096859604120255
step: 50, loss: 0.007447713054716587
step: 60, loss: 0.007550954818725586
step: 70, loss: 0.03160368278622627
step: 80, loss: 0.051487620919942856
step: 90, loss: 0.01816645823419094
step: 100, loss: 0.0017892951145768166
step: 110, loss: 0.014257506467401981
```

```
step: 120, loss: 0.003895576111972332
step: 130, loss: 0.0004540827067103237
step: 140, loss: 0.0008876165375113487
step: 150, loss: 0.001971539808437228
step: 160, loss: 0.005230551119893789
step: 170, loss: 0.011994075961411
step: 180, loss: 0.054527755826711655
step: 190, loss: 0.009965937584638596
step: 200, loss: 0.00038503811811096966
step: 210, loss: 0.00014313467545434833
step: 220, loss: 0.0001873793953564018
step: 230, loss: 0.00017895903147291392
/usr/local/lib/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=0.987598647125141, f1=0.9819004524886877,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002598507853690535
step: 10, loss: 0.00015220927889458835
step: 20, loss: 0.00015741396055091172
step: 30, loss: 0.0059108613058924675
step: 40, loss: 0.00083031126996502228
step: 50, loss: 0.0002043665444944054
step: 60, loss: 0.0021697983611375093
step: 70, loss: 0.00011499928223202005
step: 80, loss: 0.000256509956670925
step: 90, loss: 0.05570722371339798
step: 100, loss: 0.14772742986679077
step: 110, loss: 0.04293370246887207
step: 120, loss: 0.08123709261417389
step: 130, loss: 0.15478765964508057
step: 140, loss: 0.00032501333043910563
step: 150, loss: 0.00025427996297366917
step: 160, loss: 0.023967044427990913
step: 170, loss: 0.03064994513988495
step: 180, loss: 0.001155468518845737
step: 190, loss: 0.00048664724454283714
step: 200, loss: 0.006192595697939396
step: 210, loss: 0.0007416167645715177
step: 220, loss: 0.0028357820119708776
step: 230, loss: 0.00034557137405499816
```

```
/usr/local/lib/python3.7/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=0.9819819819819819, f1=0.9774774774774775,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000230037810979411
step: 10, loss: 0.0010801036842167377
step: 20, loss: 0.00027590125682763755
step: 30, loss: 0.00464500579982996
step: 40, loss: 0.0036224413197487593
step: 50, loss: 0.002137255622074008
step: 60, loss: 0.0323198139667511
step: 70, loss: 0.0004894478479400277
step: 80, loss: 0.0007424642099067569
step: 90, loss: 0.00021472095977514982
step: 100, loss: 0.00036631146213039756
step: 110, loss: 0.057074688374996185
step: 120, loss: 0.0008772448054514825
step: 130, loss: 0.012738361954689026
step: 140, loss: 0.0026061683893203735
step: 150, loss: 0.0002708060492295772
step: 160, loss: 0.016785085201263428
step: 170, loss: 8.030028402572498e-05
step: 180, loss: 0.003133243415504694
step: 190, loss: 0.0001408162497682497
step: 200, loss: 0.0006305807619355619
step: 210, loss: 0.00034316006349399686
step: 220, loss: 0.00017630349611863494
step: 230, loss: 0.012326297350227833
/usr/local/lib/python3.7/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=0.9744160177975528, f1=0.9799107142857142,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14811834692955017
step: 10, loss: 0.00010379358718637377
step: 20, loss: 0.00013376314018387347
step: 30, loss: 0.0011252070544287562
step: 40, loss: 0.005352144129574299
step: 50, loss: 0.0001523020036984235
step: 60, loss: 0.0004885124508291483
step: 70, loss: 0.00028464040951803327
```

```
step: 80, loss: 3.948593439417891e-05
step: 90, loss: 0.00022136338520795107
step: 100, loss: 0.00036950132925994694
step: 110, loss: 0.00010303337330697104
step: 120, loss: 0.0003084145428147167
step: 130, loss: 0.0005954751977697015
step: 140, loss: 0.00010916525934590027
step: 150, loss: 0.000596197322010994
step: 160, loss: 0.00015902466839179397
step: 170, loss: 0.0001189625181723386
step: 180, loss: 0.00041651277570053935
step: 190, loss: 0.0006277101347222924
step: 200, loss: 0.005100029520690441
step: 210, loss: 9.714616317069158e-05
step: 220, loss: 0.02140699326992035
step: 230, loss: 0.009861903265118599
/usr/local/lib/python3.7/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=0.9841269841269841, f1=0.9796380090497738,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011558353900909424
step: 10, loss: 0.053248923271894455
step: 20, loss: 0.001673318212851882
step: 30, loss: 9.559251338941976e-05
step: 40, loss: 0.005930169485509396
step: 50, loss: 0.001099880551919341
step: 60, loss: 0.00037197553319856524
step: 70, loss: 0.0009337427327409387
step: 80, loss: 0.002208811929449439
step: 90, loss: 0.0002208854421041906
step: 100, loss: 0.00015646881365682893
step: 110, loss: 0.000343429041095078
step: 120, loss: 0.00018393339996691793
step: 130, loss: 0.031506992876529694
step: 140, loss: 0.0036461909767240286
step: 150, loss: 0.00023488073202315718
step: 160, loss: 0.00017813843442127109
step: 170, loss: 0.001651494181714952
step: 180, loss: 0.08600926399230957
step: 190, loss: 0.002366876695305109
step: 200, loss: 0.001266211038455367
step: 210, loss: 0.0007920893840491772
step: 220, loss: 0.00041966052958741784
step: 230, loss: 0.07943630963563919
```

```
/usr/local/lib/python3.7/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=0.9796380090497738,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00040483727934770286
step: 10, loss: 0.00045760211651213467
step: 20, loss: 0.0020857793278992176
step: 30, loss: 9.967546066036448e-05
step: 40, loss: 8.992804941954091e-05
step: 50, loss: 0.0010885420488193631
step: 60, loss: 0.00011961573181906715
step: 70, loss: 0.00019319637794978917
step: 80, loss: 0.003151233773678541
step: 90, loss: 0.02429199405014515
step: 100, loss: 0.0001558945805300027
step: 110, loss: 0.0008499719551764429
step: 120, loss: 0.01989022269845009
step: 130, loss: 0.0002770903811324388
step: 140, loss: 8.367143891518936e-05
step: 150, loss: 6.81719757267274e-05
step: 160, loss: 0.00014631933299824595
step: 170, loss: 0.0001294393587158993
step: 180, loss: 0.0005015673232264817
step: 190, loss: 9.692905587144196e-05
step: 200, loss: 0.0001619108661543578
step: 210, loss: 0.00012587648234330118
step: 220, loss: 0.0457153245806694
step: 230, loss: 0.0015565184876322746
/usr/local/lib/python3.7/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=0.9831271091113611, f1=0.9808773903262092,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013454857980832458
step: 10, loss: 0.00011246653593843803
step: 20, loss: 0.0001350859965896234
step: 30, loss: 0.00022471109696198255
step: 40, loss: 8.336139580933377e-05
step: 50, loss: 0.011122792027890682
step: 60, loss: 0.00460242573171854
step: 70, loss: 0.002269066870212555
```

```
step: 80, loss: 0.0002989872300531715
step: 90, loss: 0.0001206442029797472
step: 100, loss: 0.00019192152831237763
step: 110, loss: 0.0002544111921451986
step: 120, loss: 0.00029956985963508487
step: 130, loss: 0.00010666340676834807
step: 140, loss: 0.002572042867541313
step: 150, loss: 7.504545646952465e-05
step: 160, loss: 0.0035912024322897196
step: 170, loss: 8.635673293611035e-05
step: 180, loss: 0.00015490395890083164
step: 190, loss: 0.0002607166243251413
step: 200, loss: 6.71548186801374e-05
step: 210, loss: 7.26901416783221e-05
step: 220, loss: 0.005822884384542704
step: 230, loss: 0.008278920315206051
/usr/local/lib/python3.7/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=0.9713631156930126, f1=0.9704545454545453,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22367221117019653
step: 10, loss: 0.0005695164436474442
step: 20, loss: 0.0005205004126764834
step: 30, loss: 0.007007822394371033
step: 40, loss: 0.000301172403851524
step: 50, loss: 0.0009646990802139044
step: 60, loss: 0.01480859611183405
step: 70, loss: 0.00020724347268696874
step: 80, loss: 0.0001728724892018363
step: 90, loss: 0.0004943989915773273
step: 100, loss: 0.0001888540864456445
step: 110, loss: 0.00012369266187306494
step: 120, loss: 0.00018309196457266808
step: 130, loss: 0.0001095619227271527
step: 140, loss: 8.904319111024961e-05
step: 150, loss: 7.405672658933327e-05
step: 160, loss: 9.866168693406507e-05
step: 170, loss: 0.0003472176322247833
step: 180, loss: 0.0002889815950766206
step: 190, loss: 0.04340669885277748
step: 200, loss: 0.00020245052292011678
step: 210, loss: 7.937081682030112e-05
step: 220, loss: 0.00010340238077333197
step: 230, loss: 9.506326750852168e-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=0.9842696629213483, f1=0.9820224719101124,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007982023525983095
step: 10, loss: 0.0008308945689350367
step: 20, loss: 0.00013479360495693982
step: 30, loss: 0.0002131281653419137
step: 40, loss: 5.8222023653797805e-05
step: 50, loss: 5.6563716498203576e-05
step: 60, loss: 4.661549974116497e-05
step: 70, loss: 6.0815291362814605e-05
step: 80, loss: 5.703339775209315e-05
step: 90, loss: 5.8016714319819584e-05
step: 100, loss: 4.5837328798370436e-05
step: 110, loss: 0.000238133710809052
step: 120, loss: 6.623115768888965e-05
step: 130, loss: 9.505372145213187e-05
step: 140, loss: 6.988425593590364e-05
step: 150, loss: 0.00014677460421808064
step: 160, loss: 0.004108796361833811
step: 170, loss: 9.847783076111227e-05
step: 180, loss: 0.00012081008026143536
step: 190, loss: 0.00010658584506018087
step: 200, loss: 0.00015204241208266467
step: 210, loss: 7.19773888704367e-05
step: 220, loss: 0.0007149226148612797
step: 230, loss: 8.574755338486284e-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=0.9853107344632768, f1=0.9819413092550789,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022618332877755165
step: 10, loss: 0.000131760083604604
step: 20, loss: 0.00010211672633886337
step: 30, loss: 0.0006847758777439594
step: 40, loss: 0.0002153342793462798
step: 50, loss: 9.385942394146696e-05
step: 60, loss: 6.003048474667594e-05
step: 70, loss: 6.601044879062101e-05
```

```
step: 80, loss: 5.6724195019342005e-05
step: 90, loss: 5.074578439234756e-05
step: 100, loss: 0.00018391227058600634
step: 110, loss: 6.015559483785182e-05
step: 120, loss: 6.962693441892043e-05
step: 130, loss: 0.00019339320715516806
step: 140, loss: 0.00018913055828306824
step: 150, loss: 8.04311566753313e-05
step: 160, loss: 5.091646744403988e-05
step: 170, loss: 7.55809887778014e-05
step: 180, loss: 0.0001543686812510714
step: 190, loss: 5.423799666459672e-05
step: 200, loss: 0.00011034883209504187
step: 210, loss: 0.00015332188922911882
step: 220, loss: 4.278848791727796e-05
step: 230, loss: 5.065633013145998e-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=0.9853438556933484, f1=0.9786276715410572,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.839705019956455e-05
step: 10, loss: 0.00017978093811969273
step: 20, loss: 0.0007775278063490987
step: 30, loss: 0.00016575230984017253
step: 40, loss: 3.242328966734931e-05
step: 50, loss: 9.448299533687532e-05
step: 60, loss: 3.291502798674628e-05
step: 70, loss: 0.0001278764830203727
step: 80, loss: 0.00019038595200981945
step: 90, loss: 0.0007376212161034346
step: 100, loss: 0.0003060697054024786
step: 110, loss: 0.00010701997234718874
step: 120, loss: 8.595500548835844e-05
step: 130, loss: 0.00012064012116752565
step: 140, loss: 5.3098257922101766e-05
step: 150, loss: 0.00014114481746219099
step: 160, loss: 6.216271867742762e-05
step: 170, loss: 0.00013061007484793663
step: 180, loss: 0.0004385205393191427
step: 190, loss: 0.0002645905187819153
step: 200, loss: 5.343092925613746e-05
step: 210, loss: 0.03885585442185402
step: 220, loss: 0.0002867743605747819
step: 230, loss: 2.4101855160552077e-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=0.9841628959276018, f1=0.9808773903262092,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.14911041338928e-05
step: 10, loss: 0.0005275494186207652
step: 20, loss: 0.0005590334767475724
step: 30, loss: 0.0001130450182245113
step: 40, loss: 0.00026049904408864677
step: 50, loss: 0.00010445087536936626
step: 60, loss: 4.032373908557929e-05
step: 70, loss: 0.00019773244275711477
step: 80, loss: 6.69444416416809e-05
step: 90, loss: 5.083467112854123e-05
step: 100, loss: 0.00020085890719201416
step: 110, loss: 7.054467278067023e-05
step: 120, loss: 8.883611735654995e-05
step: 130, loss: 0.0004151244356762618
step: 140, loss: 0.0006138837779872119
step: 150, loss: 0.005896606482565403
step: 160, loss: 8.381083171116188e-05
step: 170, loss: 4.774918852490373e-05
step: 180, loss: 6.241511437110603e-05
step: 190, loss: 0.0009156551095657051
step: 200, loss: 0.00021917428239248693
step: 210, loss: 0.00033153867116197944
step: 220, loss: 8.316699677379802e-05
step: 230, loss: 0.00014443877444136888
/usr/local/lib/python3.7/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=0.9853107344632768, f1=0.9786276715410572,
best_f1=0.9819004524886877
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 [00:10, 234.53it/s]
load_f1 = 0.9864559819413092
real_f1 = 0.9864864864864865
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, 250.09it/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/
```

```
Downloading: 100% 442/442 [00:00<00:00, 357kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 255kB/s]
Downloading: 100% 268M/268M [00:03<00:00, 71.9MB/s]
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: 0.7968354225158691
step: 10, loss: 0.44304507970809937
step: 20, loss: 0.48136988282203674
step: 30, loss: 0.3968265652656555
step: 40, loss: 0.2876893877983093
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 50, loss: 0.17957185208797455
step: 60, loss: 0.26900696754455566
step: 70, loss: 0.16058260202407837
step: 80, loss: 0.1799752563238144
step: 90, loss: 0.15086384117603302
step: 100, loss: 0.30457451939582825
step: 110, loss: 0.28727781772613525
step: 120, loss: 0.06308749318122864
step: 130, loss: 0.024222280830144882
step: 140, loss: 0.2094602733850479
step: 150, loss: 0.06084046512842178
step: 160, loss: 0.13779383897781372
step: 170, loss: 0.279487282037735
step: 180, loss: 0.05473006144165993
step: 190, loss: 0.011847392655909061
step: 200, loss: 0.18682602047920227
step: 210, loss: 0.10071155428886414
step: 220, loss: 0.09744393080472946
step: 230, loss: 0.16912877559661865
step: 240, loss: 0.10461672395467758
step: 250, loss: 0.061468642204999924
step: 260, loss: 0.03097568079829216
step: 270, loss: 0.034161899238824844
step: 280, loss: 0.12244072556495667
step: 290, loss: 0.07123703509569168
step: 300, loss: 0.10155724734067917
step: 310, loss: 0.06906306743621826
step: 320, loss: 0.11787798255681992
step: 330, loss: 0.0976455956697464
step: 340, loss: 0.2447877675294876
step: 350, loss: 0.07941299676895142
step: 360, loss: 0.09592771530151367
step: 370, loss: 0.14185386896133423
step: 380, loss: 0.11475177854299545
step: 390, loss: 0.22082394361495972
step: 400, loss: 0.01721854694187641
```

```
step: 410, loss: 0.08315247297286987
step: 420, loss: 0.060538750141859055
step: 430, loss: 0.037440307438373566
step: 440, loss: 0.16903969645500183
step: 450, loss: 0.008483356796205044
step: 460, loss: 0.05213198810815811
step: 470, loss: 0.22235515713691711
step: 480, loss: 0.16705025732517242
step: 490, loss: 0.03374022990465164
step: 500, loss: 0.01868225261569023
step: 510, loss: 0.10717910528182983
step: 520, loss: 0.11268239468336105
step: 530, loss: 0.09573248028755188
/usr/local/lib/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=0.9148148148148147, f1=0.908414690841469,
best_f1=0.908414690841469
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17102740705013275
step: 10, loss: 0.14139620959758759
step: 20, loss: 0.13631965219974518
step: 30, loss: 0.09590708464384079
step: 40, loss: 0.07827848941087723
step: 50, loss: 0.11584366112947464
step: 60, loss: 0.10633717477321625
step: 70, loss: 0.18019823729991913
step: 80, loss: 0.03237107768654823
step: 90, loss: 0.10358940809965134
step: 100, loss: 0.1957552134990692
step: 110, loss: 0.029023772105574608
step: 120, loss: 0.17208507657051086
step: 130, loss: 0.0649079903960228
step: 140, loss: 0.06634864956140518
step: 150, loss: 0.02331514097750187
step: 160, loss: 0.09276401996612549
step: 170, loss: 0.09444481134414673
step: 180, loss: 0.010465537197887897
step: 190, loss: 0.07736855000257492
step: 200, loss: 0.05487672984600067
step: 210, loss: 0.04165758937597275
step: 220, loss: 0.11194197088479996
```

```
step: 230, loss: 0.050307951867580414
step: 240, loss: 0.17730773985385895
step: 250, loss: 0.041493162512779236
step: 260, loss: 0.05442693084478378
step: 270, loss: 0.10074148327112198
step: 280, loss: 0.21641936898231506
step: 290, loss: 0.05512848123908043
step: 300, loss: 0.03735052049160004
step: 310, loss: 0.07708937674760818
step: 320, loss: 0.1002558022737503
step: 330, loss: 0.08896391093730927
step: 340, loss: 0.08724267035722733
step: 350, loss: 0.052640192210674286
step: 360, loss: 0.03799691051244736
step: 370, loss: 0.04777926951646805
step: 380, loss: 0.06132218986749649
step: 390, loss: 0.03628234192728996
step: 400, loss: 0.049611277878284454
step: 410, loss: 0.007837832905352116
step: 420, loss: 0.04589012637734413
step: 430, loss: 0.021102892234921455
step: 440, loss: 0.018791083246469498
step: 450, loss: 0.023288922384381294
step: 460, loss: 0.1839255392551422
step: 470, loss: 0.09519235789775848
step: 480, loss: 0.2135031372308731
step: 490, loss: 0.08533160388469696
step: 500, loss: 0.02368772402405739
step: 510, loss: 0.07700824737548828
step: 520, loss: 0.04136170074343681
step: 530, loss: 0.14393682777881622
/usr/local/lib/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=0.9252506836827713, f1=0.9215867158671587,
best_f1=0.9215867158671587
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.048550404608249664
step: 10, loss: 0.0740816667675972
step: 20, loss: 0.1615346372127533
step: 30, loss: 0.19151811301708221
step: 40, loss: 0.04393600672483444
```

```
step: 50, loss: 0.057276830077171326
step: 60, loss: 0.01208544336259365
step: 70, loss: 0.008608841337263584
step: 80, loss: 0.027646800503134727
step: 90, loss: 0.03286103159189224
step: 100, loss: 0.008667361922562122
step: 110, loss: 0.01675771363079548
step: 120, loss: 0.027203015983104706
step: 130, loss: 0.017328279092907906
step: 140, loss: 0.018846875056624413
step: 150, loss: 0.0113805141299963
step: 160, loss: 0.004561982583254576
step: 170, loss: 0.010682755149900913
step: 180, loss: 0.004785060882568359
step: 190, loss: 0.029731713235378265
step: 200, loss: 0.022741857916116714
step: 210, loss: 0.0716443806886673
step: 220, loss: 0.0320950485765934
step: 230, loss: 0.03846540302038193
step: 240, loss: 0.038997918367385864
step: 250, loss: 0.046033747494220734
step: 260, loss: 0.005731913726776838
step: 270, loss: 0.0018544200574979186
step: 280, loss: 0.011857268400490284
step: 290, loss: 0.0901707261800766
step: 300, loss: 0.10161896049976349
step: 310, loss: 0.04312385246157646
step: 320, loss: 0.1608390361070633
step: 330, loss: 0.0022275832016021013
step: 340, loss: 0.04338511452078819
step: 350, loss: 0.013584607280790806
step: 360, loss: 0.007591349072754383
step: 370, loss: 0.011814558878540993
step: 380, loss: 0.006497617810964584
step: 390, loss: 0.04876234382390976
step: 400, loss: 0.03428361564874649
step: 410, loss: 0.02689996361732483
step: 420, loss: 0.06372881680727005
step: 430, loss: 0.05587342754006386
step: 440, loss: 0.05761958658695221
step: 450, loss: 0.10316043347120285
step: 460, loss: 0.08627672493457794
step: 470, loss: 0.02776058204472065
step: 480, loss: 0.005009936168789864
step: 490, loss: 0.08107520639896393
step: 500, loss: 0.09829043596982956
step: 510, loss: 0.017689237371087074
step: 520, loss: 0.013879078440368176
```

```
step: 530, loss: 0.011802812106907368
/usr/local/lib/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=0.929683813119396, f1=0.9155937052932761,
best_f1=0.9155937052932761
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011280394159257412
step: 10, loss: 0.011747512966394424
step: 20, loss: 0.011792528443038464
step: 30, loss: 0.06670349836349487
step: 40, loss: 0.04071987420320511
step: 50, loss: 0.02863450162112713
step: 60, loss: 0.008504031226038933
step: 70, loss: 0.00048048325697891414
step: 80, loss: 0.07061765342950821
step: 90, loss: 0.014926665462553501
step: 100, loss: 0.0264894962310791
step: 110, loss: 0.0005815106560476124
step: 120, loss: 0.00033815018832683563
step: 130, loss: 0.019071724265813828
step: 140, loss: 0.2808586061000824
step: 150, loss: 0.018312670290470123
step: 160, loss: 0.014061233960092068
step: 170, loss: 0.005527842324227095
step: 180, loss: 0.02464132010936737
step: 190, loss: 0.12258222699165344
step: 200, loss: 0.0027017162647098303
step: 210, loss: 0.029915286228060722
step: 220, loss: 0.04547489434480667
step: 230, loss: 0.24162830412387848
step: 240, loss: 0.004778516944497824
step: 250, loss: 0.0031850284431129694
step: 260, loss: 0.13984401524066925
step: 270, loss: 0.16598616540431976
step: 280, loss: 0.006216213572770357
step: 290, loss: 0.09606269001960754
step: 300, loss: 0.003345467848703265
step: 310, loss: 0.0675409659743309
step: 320, loss: 0.0193462073802948
step: 330, loss: 0.037743814289569855
step: 340, loss: 0.06453754007816315
```

```
step: 350, loss: 0.0014020424569025636
step: 360, loss: 0.04227360710501671
step: 370, loss: 0.037997566163539886
step: 380, loss: 0.010226823389530182
step: 390, loss: 0.011790823191404343
step: 400, loss: 0.014641688205301762
step: 410, loss: 0.06244262680411339
step: 420, loss: 0.009703165851533413
step: 430, loss: 0.012638965621590614
step: 440, loss: 0.1694469004869461
step: 450, loss: 0.020564882084727287
step: 460, loss: 0.0012056288542225957
step: 470, loss: 0.03847520053386688
step: 480, loss: 0.002301652217283845
step: 490, loss: 0.01369383092969656
step: 500, loss: 0.016113372519612312
step: 510, loss: 0.01973041146993637
step: 520, loss: 0.08550753444433212
step: 530, loss: 0.002864051377400756
/usr/local/lib/python3.7/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=0.9280742459396751, f1=0.9141248240262787,
best_f1=0.9155937052932761
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004346649162471294
step: 10, loss: 0.045563071966171265
step: 20, loss: 0.10296762734651566
step: 30, loss: 0.011067538522183895
step: 40, loss: 0.034921955317258835
step: 50, loss: 0.004729543812572956
step: 60, loss: 0.001117688836529851
step: 70, loss: 0.0016830030363053083
step: 80, loss: 0.008722328580915928
step: 90, loss: 0.10116759687662125
step: 100, loss: 0.0051616099663078785
step: 110, loss: 0.012756673619151115
step: 120, loss: 0.008985953405499458
step: 130, loss: 0.0020778661128133535
step: 140, loss: 0.0013576109195128083
step: 150, loss: 0.00452836649492383
step: 160, loss: 0.011837242171168327
step: 170, loss: 0.023751109838485718
step: 180, loss: 0.0653298944234848
step: 190, loss: 0.0050149941816926
step: 200, loss: 0.04838060215115547
```

```
step: 210, loss: 0.005914926528930664
step: 220, loss: 0.0013075286988168955
step: 230, loss: 0.0018650070996955037
step: 240, loss: 0.0502932108938694
step: 250, loss: 0.00033326499396935105
step: 260, loss: 0.00029946694849058986
step: 270, loss: 0.034939855337142944
step: 280, loss: 0.03771857172250748
step: 290, loss: 0.0011409127619117498
step: 300, loss: 0.0974142923951149
step: 310, loss: 0.0017961720004677773
step: 320, loss: 0.0016428162343800068
step: 330, loss: 0.018557660281658173
step: 340, loss: 0.0017256083665415645
step: 350, loss: 0.023376306518912315
step: 360, loss: 0.0370015986263752
step: 370, loss: 0.01334522943943739
step: 380, loss: 0.024411410093307495
step: 390, loss: 0.001434290548786521
step: 400, loss: 0.07986916601657867
step: 410, loss: 0.0012666040565818548
step: 420, loss: 0.0009231339790858328
step: 430, loss: 0.04303378239274025
step: 440, loss: 0.012951353564858437
step: 450, loss: 0.025495290756225586
step: 460, loss: 0.0029166426975280046
step: 470, loss: 0.0018809556495398283
step: 480, loss: 0.016337888315320015
step: 490, loss: 0.006808932404965162
step: 500, loss: 0.0030755929183214903
step: 510, loss: 0.007189322728663683
step: 520, loss: 0.0010188001906499267
step: 530, loss: 0.0037040580064058304
/usr/local/lib/python3.7/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=0.9232914923291492, f1=0.9186915887850468,
best_f1=0.9155937052932761
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005863464903086424
step: 10, loss: 0.0013747074408456683
step: 20, loss: 0.002959738252684474
step: 30, loss: 0.0009295229101553559
step: 40, loss: 0.004386994522064924
step: 50, loss: 0.004091565031558275
step: 60, loss: 0.000307201175019145
```

```
step: 70, loss: 0.0011725964723154902
step: 80, loss: 0.13224337995052338
step: 90, loss: 0.0007721259025856853
step: 100, loss: 0.04715043306350708
step: 110, loss: 0.0010722960578277707
step: 120, loss: 0.006470688618719578
step: 130, loss: 0.003638050053268671
step: 140, loss: 0.0111769437789917
step: 150, loss: 0.001440006191842258
step: 160, loss: 0.0006959815509617329
step: 170, loss: 0.004070380236953497
step: 180, loss: 0.002777640474960208
step: 190, loss: 0.20140469074249268
step: 200, loss: 0.00037275697104632854
step: 210, loss: 0.00033725507091730833
step: 220, loss: 0.0007233446813188493
step: 230, loss: 0.00013579650840256363
step: 240, loss: 0.00851423293352127
step: 250, loss: 0.0015674418536946177
step: 260, loss: 0.00042329321149736643
step: 270, loss: 0.0011823985259979963
step: 280, loss: 0.007262322120368481
step: 290, loss: 0.0006095776916481555
step: 300, loss: 0.0011694671120494604
step: 310, loss: 0.011513743549585342
step: 320, loss: 0.002885140012949705
step: 330, loss: 0.05690143257379532
step: 340, loss: 0.007925482466816902
step: 350, loss: 0.005362948402762413
step: 360, loss: 0.00013250349729787558
step: 370, loss: 0.0010453058639541268
step: 380, loss: 0.0005276046576909721
step: 390, loss: 0.0035119440872222185
step: 400, loss: 0.000133686262415722
step: 410, loss: 0.001046499703079462
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 420, loss: 0.19963984191417694
step: 430, loss: 8.717968012206256e-05
step: 440, loss: 0.03176097199320793
step: 450, loss: 0.023550422862172127
step: 460, loss: 0.07798872888088226
step: 470, loss: 0.11667679250240326
step: 480, loss: 0.012759391218423843
step: 490, loss: 0.0013360442826524377
step: 500, loss: 0.005882470402866602
step: 510, loss: 0.0012451306683942676
step: 520, loss: 0.0003967273223679513
step: 530, loss: 0.006406531669199467
```

```
/usr/local/lib/python3.7/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=0.9084041548630784, f1=0.8917501192179303,
best_f1=0.9155937052932761
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07658545672893524
step: 10, loss: 0.012790342792868614
step: 20, loss: 0.011559785343706608
step: 30, loss: 0.0011587155750021338
step: 40, loss: 0.00031462209881283343
step: 50, loss: 0.0021220482885837555
step: 60, loss: 0.0017528630560263991
step: 70, loss: 0.0035987510345876217
step: 80, loss: 0.00031390716321766376
step: 90, loss: 0.01961822435259819
step: 100, loss: 0.0009531056857667863
step: 110, loss: 0.03385188430547714
step: 120, loss: 0.00012942765897605568
step: 130, loss: 0.0006880330038256943
step: 140, loss: 0.002314346143975854
step: 150, loss: 0.00016604605480097234
step: 160, loss: 0.0003563326317816973
step: 170, loss: 5.086588134872727e-05
step: 180, loss: 0.0027444041334092617
step: 190, loss: 4.515264299698174e-05
step: 200, loss: 0.0006385224405676126
step: 210, loss: 0.004206468351185322
step: 220, loss: 0.000504291441757232
step: 230, loss: 0.0006141621270217001
step: 240, loss: 0.007111991755664349
step: 250, loss: 7.355460547842085e-05
step: 260, loss: 0.016422532498836517
step: 270, loss: 0.00011751452984753996
step: 280, loss: 0.07485263049602509
step: 290, loss: 0.0006930292584002018
step: 300, loss: 0.00028006607317365706
step: 310, loss: 0.002076852833852172
step: 320, loss: 0.038950640708208084
step: 330, loss: 4.409862231113948e-05
step: 340, loss: 0.02475927770137787
step: 350, loss: 0.000496997032314539
step: 360, loss: 0.0028872406110167503
step: 370, loss: 0.0004184468707535416
step: 380, loss: 0.0004246430180501193
step: 390, loss: 8.684198837727308e-05
```

```
step: 400, loss: 0.0031984401866793633
step: 410, loss: 0.0005605798796750605
step: 420, loss: 0.0002087416942231357
step: 430, loss: 9.331451292382553e-05
step: 440, loss: 0.007479261141270399
step: 450, loss: 0.0010642242850735784
step: 460, loss: 0.0032991403713822365
step: 470, loss: 0.012441374361515045
step: 480, loss: 0.0015584176871925592
step: 490, loss: 6.53402748866938e-05
step: 500, loss: 0.00012138303281972185
step: 510, loss: 0.00749979866668582
step: 520, loss: 0.00043725845171138644
step: 530, loss: 0.0006107548833824694
/usr/local/lib/python3.7/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=0.9197355996222852, f1=0.9001919385796545,
best_f1=0.9155937052932761
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008888945914804935
step: 10, loss: 0.0005020964308641851
step: 20, loss: 0.00041320090531371534
step: 30, loss: 0.010865116491913795
step: 40, loss: 0.0030067607294768095
step: 50, loss: 0.0003417053958401084
step: 60, loss: 0.0002664256899151951
step: 70, loss: 0.001285766251385212
step: 80, loss: 0.0007300337310880423
step: 90, loss: 5.478774619405158e-05
step: 100, loss: 0.07383977621793747
step: 110, loss: 0.0016193357296288013
step: 120, loss: 9.097060683416203e-05
step: 130, loss: 0.03052004426717758 2
step: 140, loss: 2.9727278160862625e-05
step: 150, loss: 0.0010463926009833813
step: 160, loss: 0.01039094664156437
step: 170, loss: 8.774151501711458e-05
step: 180, loss: 0.0005463623092509806
step: 190, loss: 0.0005401002708822489
step: 200, loss: 0.0039043861906975508
step: 210, loss: 0.001931625185534358
step: 220, loss: 0.00024714329629205 17
step: 230, loss: 0.00049143738579005
step: 240, loss: 0.001349331345409155
step: 250, loss: 0.00030368976877070963
```

```
step: 260, loss: 0.023551587015390396
step: 270, loss: 9.707700519356877e-05
step: 280, loss: 0.0009790033800527453
step: 290, loss: 0.0010577647481113672
step: 300, loss: 7.733614620519802e-05
step: 310, loss: 0.032982297241687775
step: 320, loss: 0.003713322337716818
step: 330, loss: 0.05090963840484619
step: 340, loss: 0.0016448599053546786
step: 350, loss: 0.2522565424442291
step: 360, loss: 0.0029747725930064917
step: 370, loss: 0.0008170395740307868
step: 380, loss: 0.036191802471876144
step: 390, loss: 0.002666166517883539
step: 400, loss: 0.06544311344623566
step: 410, loss: 0.0002809628495015204
step: 420, loss: 0.00013512131408788264
step: 430, loss: 0.001566563150845468
step: 440, loss: 0.0025934234727174044
step: 450, loss: 0.001199399004690349
step: 460, loss: 0.00041368897655047476
step: 470, loss: 6.492374814115465e-05
step: 480, loss: 6.010836295899935e-05
step: 490, loss: 0.0015401395503431559
step: 500, loss: 0.0002239713357994333
step: 510, loss: 0.00011207682837266475
step: 520, loss: 0.0006021138979122043
step: 530, loss: 5.085305019747466e-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 8: dev_f1=0.9307583608101743, f1=0.9148834997622444,
best_f1=0.9148834997622444
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017839187057688832
step: 10, loss: 0.0005468434537760913
step: 20, loss: 9.569185203872621e-05
step: 30, loss: 7.063525845296681e-05
step: 40, loss: 0.003948703408241272
step: 50, loss: 0.00014541005657520145
step: 60, loss: 8.973816147772595e-05
step: 70, loss: 0.13341949880123138
```

```
step: 80, loss: 0.0003556316369213164
step: 90, loss: 0.0008448614389635623
step: 100, loss: 0.001228039851412177
step: 110, loss: 9.665051766205579e-05
step: 120, loss: 7.581575482618064e-05
step: 130, loss: 0.0023390264250338078
step: 140, loss: 9.040913573699072e-05
step: 150, loss: 7.741117588011548e-05
step: 160, loss: 0.00027614127611741424
step: 170, loss: 0.0797727033495903
step: 180, loss: 0.0001412939018337056
step: 190, loss: 0.00025930965784937143
step: 200, loss: 0.0013389443047344685
step: 210, loss: 3.83318547392264e-05
step: 220, loss: 6.20696009718813e-05
step: 230, loss: 0.0003385498421266675
step: 240, loss: 0.002576257335022092
step: 250, loss: 0.00031923511414788663
step: 260, loss: 8.470850298181176e-05
step: 270, loss: 0.0032469904981553555
step: 280, loss: 9.014371607918292e-05
step: 290, loss: 0.015708990395069122
step: 300, loss: 9.088224760489538e-05
step: 310, loss: 3.5482345992932096e-05
step: 320, loss: 0.012447206303477287
step: 330, loss: 7.464316877303645e-05
step: 340, loss: 5.2142931963317096e-05
step: 350, loss: 4.419137621880509e-05
step: 360, loss: 0.008911955170333385
step: 370, loss: 0.00016672494530212134
step: 380, loss: 0.00011570648348424584
step: 390, loss: 7.579071825603023e-05
step: 400, loss: 0.0008286124793812633
step: 410, loss: 0.00028432844555936754
step: 420, loss: 0.000985668390057981
step: 430, loss: 8.871091267792508e-05
step: 440, loss: 0.007703524082899094
step: 450, loss: 0.0003992928541265428
step: 460, loss: 0.0016636423533782363
step: 470, loss: 0.0005401261150836945
step: 480, loss: 0.00590873695909977
step: 490, loss: 0.00014940882101655006
step: 500, loss: 0.004742654040455818
step: 510, loss: 0.00046824588207527995
step: 520, loss: 0.00019758699636440724
step: 530, loss: 0.00016099376080092043
/usr/local/lib/python3.7/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=0.9263256687001408, f1=0.9130434782608695,
best_f1=0.9148834997622444
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001549103035358712
step: 10, loss: 3.6518005799734965e-05
step: 20, loss: 5.675013017025776e-05
step: 30, loss: 6.18453195784241e-05
step: 40, loss: 8.682592306286097e-05
step: 50, loss: 3.7013574910815805e-05
step: 60, loss: 0.0002475691435392946
step: 70, loss: 0.009739778935909271
step: 80, loss: 0.010684844106435776
step: 90, loss: 0.0001296540576731786
step: 100, loss: 4.951618393533863e-05
step: 110, loss: 0.00012846746540162712
step: 120, loss: 0.00011961517157033086
step: 130, loss: 5.535899254027754e-05
step: 140, loss: 0.00045405025593936443
step: 150, loss: 0.00032707874197512865
step: 160, loss: 0.0001678965927567333
step: 170, loss: 6.142577331047505e-05
step: 180, loss: 0.0020129128824919462
step: 190, loss: 5.0961196393473074e-05
step: 200, loss: 0.0001140629465226084
step: 210, loss: 4.404898936627433e-05
step: 220, loss: 0.004497154615819454
step: 230, loss: 0.00030662454082630575
step: 240, loss: 4.4821525079896674e-05
step: 250, loss: 0.0002440370008116588
step: 260, loss: 0.00014906896103639156
step: 270, loss: 3.5146305890521035e-05
step: 280, loss: 0.00011157553672092035
step: 290, loss: 5.1223309128545225e-05
step: 300, loss: 6.81057499605231e-05
step: 310, loss: 3.619716881075874e-05
step: 320, loss: 0.0002991599903907627
step: 330, loss: 9.253061580238864e-05
step: 340, loss: 5.959881309536286e-05
step: 350, loss: 3.635411849245429e-05
step: 360, loss: 0.004403538536280394
step: 370, loss: 0.0010991966119036078
step: 380, loss: 3.113503771601245e-05
step: 390, loss: 5.041117765358649e-05
step: 400, loss: 4.330007141106762e-05
step: 410, loss: 0.0006703975377604365
```

```
step: 420, loss: 0.0007281620055437088
step: 430, loss: 7.80232367105782e-05
step: 440, loss: 3.760917024919763e-05
step: 450, loss: 0.0007817979203537107
step: 460, loss: 0.0015296116471290588
step: 470, loss: 3.60032390744891e-05
step: 480, loss: 0.00011341759818606079
step: 490, loss: 0.04714745655655861
step: 500, loss: 0.0022884916979819536
step: 510, loss: 9.12520190468058e-05
step: 520, loss: 0.000192952313227579
step: 530, loss: 6.04686101723928e-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=0.9299242424242424, f1=0.914448669201521,
best_f1=0.9148834997622444
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.7672201617388055e-05
step: 10, loss: 0.0006312247132882476
step: 20, loss: 5.4648662626277655e-05
step: 30, loss: 0.0015652960864827037
step: 40, loss: 0.0002986536710523069
step: 50, loss: 0.0006051330710761249
step: 60, loss: 0.0006552593549713492
step: 70, loss: 0.00014036351058166474
step: 80, loss: 5.6598630180815235e-05
step: 90, loss: 3.1969204428605735e-05
step: 100, loss: 2.6862360755330883e-05
step: 110, loss: 3.834246308542788e-05
step: 120, loss: 0.00028975508757866692
step: 130, loss: 5.1794831961160526e-05
step: 140, loss: 4.211386476526968e-05
step: 150, loss: 3.279630618635565e-05
step: 160, loss: 0.00041233000229112804
step: 170, loss: 0.00011484744027256966
step: 180, loss: 2.9961720429128036e-05
step: 190, loss: 0.03158640116453171
step: 200, loss: 0.0003637500340119004
step: 210, loss: 0.0002765146200545132
step: 220, loss: 0.00023869439610280097
step: 230, loss: 0.00043171169818378985
step: 240, loss: 3.8966405554674566e-05
step: 250, loss: 0.0009200782515108585
step: 260, loss: 0.0016913347644731402
step: 270, loss: 0.0007981918170116842
```

```
step: 280, loss: 0.0015805161092430353
step: 290, loss: 0.0028161394875496626
step: 300, loss: 0.0017612538067623973
step: 310, loss: 0.013806439004838467
step: 320, loss: 0.0029079397208988667
step: 330, loss: 0.0010502770310267806
step: 340, loss: 3.285886123194359e-05
step: 350, loss: 0.0009959358721971512
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 360, loss: 0.055339131504297256
step: 370, loss: 3.803459549089894e-05
step: 380, loss: 9.712793689686805e-05
step: 390, loss: 0.026497893035411835
step: 400, loss: 2.895174839068204e-05
step: 410, loss: 0.00010151702008442953
step: 420, loss: 0.0003535887517500669
step: 430, loss: 0.00012771067849826068
step: 440, loss: 0.00035000918433070183
step: 450, loss: 0.0002515910309739411
step: 460, loss: 0.010605684481561184
step: 470, loss: 0.00130856700707227
step: 480, loss: 6.722364923916757e-05
step: 490, loss: 0.0013984511606395245
step: 500, loss: 3.500885213725269e-05
step: 510, loss: 4.6900444431230426e-05
step: 520, loss: 2.2381203962140717e-05
step: 530, loss: 2.425483580736909e-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=0.9307021569527306, f1=0.9195402298850575,
best_f1=0.9148834997622444
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.510071686236188e-05
step: 10, loss: 9.01838720892556e-05
step: 20, loss: 2.295489866810385e-05
step: 30, loss: 8.390226867049932e-05
step: 40, loss: 0.02107659913599491
step: 50, loss: 0.00013644943828694522
step: 60, loss: 0.0002346440451219678
step: 70, loss: 0.0008223371696658432
step: 80, loss: 0.0017293673008680344
step: 90, loss: 0.00020792636496480554
step: 100, loss: 8.82579552126117e-05
step: 110, loss: 9.010210487758741e-05
step: 120, loss: 4.808911762665957e-05
```

```
step: 130, loss: 4.674077717936598e-05
step: 140, loss: 2.802839117066469e-05
step: 150, loss: 9.017902630148456e-05
step: 160, loss: 6.198922346811742e-05
step: 170, loss: 0.00033794899354688823
step: 180, loss: 8.234556298702955e-05
step: 190, loss: 1.8521954189054668e-05
step: 200, loss: 0.0013284290907904506
step: 210, loss: 5.271115151117556e-05
step: 220, loss: 0.0028127485420554876
step: 230, loss: 0.011498669162392616
step: 240, loss: 6.989840767346323e-05
step: 250, loss: 3.8665413740091026e-05
step: 260, loss: 0.00030626857187598944
step: 270, loss: 4.0507704397896305e-05
step: 280, loss: 2.6244002583553083e-05
step: 290, loss: 0.0013679488329216838
step: 300, loss: 0.00020835339091718197
step: 310, loss: 0.00025775760877877474
step: 320, loss: 0.0007754080579616129
step: 330, loss: 2.407592546660453e-05
step: 340, loss: 0.00012949416122864932
step: 350, loss: 0.008428708650171757
step: 360, loss: 0.00018642444047145545
step: 370, loss: 4.34812500316184e-05
step: 380, loss: 0.0001960174849955365
step: 390, loss: 3.744419154827483e-05
step: 400, loss: 0.00021463325538206846
step: 410, loss: 6.261985254241154e-05
step: 420, loss: 0.0009872176451608539
step: 430, loss: 0.0023380271159112453
step: 440, loss: 0.0012961108004674315
step: 450, loss: 0.000679153308738023
step: 460, loss: 0.00018500554142519832
step: 470, loss: 6.371689960360527e-05
step: 480, loss: 3.972974081989378e-05
step: 490, loss: 3.7467751099029556e-05
step: 500, loss: 0.0053944336250424385
step: 510, loss: 3.9127560739871114e-05
step: 520, loss: 2.4328857762156986e-05
step: 530, loss: 0.03956637158989906
/usr/local/lib/python3.7/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=0.9271523178807948, f1=0.9086561453849833,
best_f1=0.9148834997622444
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.78010154236108e-05
step: 10, loss: 0.00023760991462040693
step: 20, loss: 0.0064911022782325745
step: 30, loss: 0.005019434727728367
step: 40, loss: 0.00014386991097126156
step: 50, loss: 0.0001608991005923599
step: 60, loss: 2.051115188805852e-05
step: 70, loss: 1.843253085098695e-05
step: 80, loss: 6.771759217372164e-05
step: 90, loss: 0.0004932297160848975
step: 100, loss: 1.7463951735408045e-05
step: 110, loss: 9.407808829564601e-05
step: 120, loss: 2.8437876608222723e-05
step: 130, loss: 0.0003247808199375868
step: 140, loss: 5.029755993746221e-05
step: 150, loss: 0.00017721077892929316
step: 160, loss: 0.0001106583877117373
step: 170, loss: 2.273868267366197e-05
step: 180, loss: 2.7007365133613348e-05
step: 190, loss: 0.010552987456321716
step: 200, loss: 0.0008644448826089501
step: 210, loss: 0.0009939632145687938
step: 220, loss: 5.415296618593857e-05
step: 230, loss: 0.00010099552309839055
step: 240, loss: 3.424503665883094e-05
step: 250, loss: 0.007687064819037914
step: 260, loss: 1.8618737158249132e-05
step: 270, loss: 2.454523382766638e-05
step: 280, loss: 4.0547169191995636e-05
step: 290, loss: 3.2002040825318545e-05
step: 300, loss: 4.13950729125645e-05
step: 310, loss: 3.038088107132353e-05
step: 320, loss: 3.745697540580295e-05
step: 330, loss: 0.000533485843334347
step: 340, loss: 1.7527299860375933e-05
step: 350, loss: 0.0037611688021570444
step: 360, loss: 2.2042215277906507e-05
step: 370, loss: 0.00015263179375324398
step: 380, loss: 0.004343599546700716
step: 390, loss: 0.0025139639619737864
step: 400, loss: 0.00023074015916790813
step: 410, loss: 0.00472091231495142
step: 420, loss: 3.5953573387814686e-05
step: 430, loss: 6.907093484187499e-05
step: 440, loss: 3.399335400899872e-05
step: 450, loss: 8.314011211041361e-05
step: 460, loss: 0.00012345828872639686
```

```
step: 470, loss: 0.01812973991036415
step: 480, loss: 0.00023320261971093714
step: 490, loss: 0.00029025584808550775
step: 500, loss: 0.0005069257458671927
step: 510, loss: 2.602796485007275e-05
step: 520, loss: 3.149556869175285e-05
step: 530, loss: 0.0002671845140866935
/usr/local/lib/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=0.931390977443609, f1=0.9174138744690892,
best_f1=0.9174138744690892
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021901998843532056
step: 10, loss: 9.297777432948351e-05
step: 20, loss: 3.955418651457876e-05
step: 30, loss: 5.929048347752541e-05
step: 40, loss: 0.00012303674884606153
step: 50, loss: 5.0336806452833116e-05
step: 60, loss: 0.0003602106589823961
step: 70, loss: 0.00022315057867672294
step: 80, loss: 6.329466850729659e-05
step: 90, loss: 0.0004645887529477477
step: 100, loss: 0.0006595324957743287
step: 110, loss: 3.9216825825860724e-05
step: 120, loss: 0.00011058857489842921
step: 130, loss: 6.545299402205274e-05
step: 140, loss: 0.03824256360530853
step: 150, loss: 0.0010598880471661687
step: 160, loss: 0.00027748718275688859
step: 170, loss: 0.00037694862112402916
step: 180, loss: 0.0238969586789608
step: 190, loss: 0.00031974216108210385
step: 200, loss: 7.566383283119649e-05
step: 210, loss: 0.0004887834074907005
step: 220, loss: 0.0020985815208405256
step: 230, loss: 0.000785109878052026
step: 240, loss: 0.00041474352474324405
step: 250, loss: 6.964613567106426e-05
step: 260, loss: 5.1308994443388656e-05
step: 270, loss: 0.0031002762261778116
step: 280, loss: 1.796311335056089e-05
```

```
step: 290, loss: 0.0018006728496402502
step: 300, loss: 2.9718188670813106e-05
step: 310, loss: 1.7199392459588125e-05
step: 320, loss: 3.1409388611791655e-05
step: 330, loss: 0.0004573274636641145
step: 340, loss: 0.005636861082166433
step: 350, loss: 0.0001264385791728273
step: 360, loss: 0.00035242034937255085
step: 370, loss: 5.217198850004934e-05
step: 380, loss: 0.0013929030392318964
step: 390, loss: 0.003469209186732769
step: 400, loss: 5.6890366977313533e-05
step: 410, loss: 0.004456947557628155
step: 420, loss: 5.541630889638327e-05
step: 430, loss: 2.1300844309735112e-05
step: 440, loss: 1.7329764887108468e-05
step: 450, loss: 1.807860826374963e-05
step: 460, loss: 6.662482337560505e-05
step: 470, loss: 6.111624679761007e-05
step: 480, loss: 2.7279167625238188e-05
step: 490, loss: 4.392392293084413e-05
step: 500, loss: 0.00014746739179827273
step: 510, loss: 0.0007132388418540359
step: 520, loss: 9.142322960542515e-05
step: 530, loss: 5.800027429359034e-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 14: dev_f1=0.9315960912052118, f1=0.9185667752442996,
best_f1=0.9185667752442996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.378887049620971e-05
step: 10, loss: 4.0530194382881746e-05
step: 20, loss: 2.3457652787328698e-05
step: 30, loss: 9.084003249881789e-05
step: 40, loss: 0.00015102254110388458
step: 50, loss: 0.00021047753398306668
step: 60, loss: 4.915601311950013e-05
step: 70, loss: 0.0004927407135255635
step: 80, loss: 2.1676951291738078e-05
step: 90, loss: 1.3921304343966767e-05
step: 100, loss: 2.156512528017629e-05
```

```
step: 110, loss: 0.03333834931254387
step: 120, loss: 0.0017486957367509604
step: 130, loss: 2.3613009034306742e-05
step: 140, loss: 1.4256567737902515e-05
step: 150, loss: 2.947255052276887e-05
step: 160, loss: 5.451845208881423e-05
step: 170, loss: 2.7048381525673904e-05
step: 180, loss: 1.705788599792868e-05
step: 190, loss: 4.306154005462304e-05
step: 200, loss: 1.8790078684105538e-05
step: 210, loss: 2.2880034521222115e-05
step: 220, loss: 3.654436659417115e-05
step: 230, loss: 0.0010395287536084652
step: 240, loss: 2.1829770048498176e-05
step: 250, loss: 2.241051333840005e-05
step: 260, loss: 2.129308631992899e-05
step: 270, loss: 0.00025525811361148953
step: 280, loss: 0.00023980261175893247
step: 290, loss: 0.0029577179811894894
step: 300, loss: 0.003645511112445593
step: 310, loss: 0.00037233909824863076
step: 320, loss: 2.9722708859480917e-05
step: 330, loss: 0.002785127842798829
step: 340, loss: 0.0003387558099348098
step: 350, loss: 3.473411197774112e-05
step: 360, loss: 0.00020910150487907230
step: 370, loss: 2.743338882282842e-05
step: 380, loss: 4.86642456962727e-05
step: 390, loss: 6.175631278892979e-05
step: 400, loss: 2.3374162992695346e-05
step: 410, loss: 8.356261241715401e-05
step: 420, loss: 8.756705210544169e-05
step: 430, loss: 2.0689762095571496e-05
step: 440, loss: 0.00011716660810634494
step: 450, loss: 0.00041389063699170947
step: 460, loss: 3.103958079009317e-05
step: 470, loss: 0.00020354673324618489
step: 480, loss: 1.9214641724829562e-05
step: 490, loss: 0.0014040335081517696
step: 500, loss: 0.0012880866415798664
step: 510, loss: 0.00017324398504570127
step: 520, loss: 1.5523124602623284e-05
step: 530, loss: 1.6137724742293358e-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=0.9310504396112911, f1=0.9161648911533117,
best_f1=0.9185667752442996
```

```
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.")
5742it [00:22, 259.15it/s]
load_f1 = 0.929245283018868
real_f1 = 0.9277279168634862
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, 262.52it/s]
```

1.1.4 Amazon-Google - Running the matcher

```python
# 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 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: 0.845454752445221
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.06196283549070358
step: 20, loss: 0.375750333070755
step: 30, loss: 0.37309128046035767
step: 40, loss: 0.5104390978813171
step: 50, loss: 0.30134883522987366
step: 60, loss: 0.3678000867366791
step: 70, loss: 0.24378950893878937
step: 80, loss: 0.29391148686408997
step: 90, loss: 0.43440595269203186
step: 100, loss: 0.19284865260124207
step: 110, loss: 0.28984466195106506
step: 120, loss: 0.2456839382648468
step: 130, loss: 0.23708254098892212
step: 140, loss: 0.29009950160980225
step: 150, loss: 0.27495354413986206
step: 160, loss: 0.23517315089702606
step: 170, loss: 0.19555574655532837
step: 180, loss: 0.2422739863395691
step: 190, loss: 0.22599202394485474
step: 200, loss: 0.21780140697956085
step: 210, loss: 0.4935702979564667
/usr/local/lib/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=0.5017667844522967, f1=0.5166051660516605,
best_f1=0.5166051660516605
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08591858297586441
step: 10, loss: 0.06820788979530334
step: 20, loss: 0.213332861661911
step: 30, loss: 0.19609546661376953
step: 40, loss: 0.0469328835606575
step: 50, loss: 0.228281632065773
step: 60, loss: 0.07944667339324951
step: 70, loss: 0.2172478288412094
step: 80, loss: 0.23855943977832794
step: 90, loss: 0.11980613321065903
step: 100, loss: 0.13941623270511627
step: 110, loss: 0.0744422897696495
step: 120, loss: 0.1889641284942627
step: 130, loss: 0.26038503646850586
step: 140, loss: 0.27867987751960754
step: 150, loss: 0.25096994638442993
step: 160, loss: 0.18558263778686523
step: 170, loss: 0.21861761808395386
step: 180, loss: 0.27422085404396057
step: 190, loss: 0.22727909684181213
step: 200, loss: 0.2191218137741089
step: 210, loss: 0.27539077401161194
/usr/local/lib/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=0.5515947467166981, f1=0.5782178217821783,
best_f1=0.5782178217821783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.29214128851890564
step: 10, loss: 0.18951302766799927
step: 20, loss: 0.2363492101430893
step: 30, loss: 0.07398749142885208
step: 40, loss: 0.19799427688121796
step: 50, loss: 0.14363744854927063
step: 60, loss: 0.22080670297145844
step: 70, loss: 0.1079237088561058
step: 80, loss: 0.06397662311792374
step: 90, loss: 0.12104634195566177
step: 100, loss: 0.04449314996600151
step: 110, loss: 0.2950102686882019
step: 120, loss: 0.10882513225078583
```

```
step: 130, loss: 0.18372711539268494
step: 140, loss: 0.3042854368686676
step: 150, loss: 0.2622920870780945
step: 160, loss: 0.1071968823671341
step: 170, loss: 0.13669942319393158
step: 180, loss: 0.12242982536554337
step: 190, loss: 0.33213019371032715
step: 200, loss: 0.112073615193367
step: 210, loss: 0.25906288623809814
/usr/local/lib/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=0.5644171779141104, f1=0.5657015590200446,
best_f1=0.5657015590200446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1722484678030014
step: 10, loss: 0.05289143696427345
step: 20, loss: 0.12074460089206696
step: 30, loss: 0.1068509966135025
step: 40, loss: 0.08661868423223495
step: 50, loss: 0.07543693482875824
step: 60, loss: 0.06061636283993721
step: 70, loss: 0.13410599529743195
step: 80, loss: 0.21153531968593597
step: 90, loss: 0.009224616922438145
step: 100, loss: 0.12128095328807831
step: 110, loss: 0.17450231313705444
step: 120, loss: 0.07387136667966843
step: 130, loss: 0.23992499709129333
step: 140, loss: 0.1506887972354889
step: 150, loss: 0.14197653532028198
step: 160, loss: 0.04353680834174156
step: 170, loss: 0.05625731870532036
step: 180, loss: 0.1155814528465271
step: 190, loss: 0.06293802708387375
step: 200, loss: 0.13236527144908905
step: 210, loss: 0.06700429320335388
/usr/local/lib/python3.7/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=0.5481798715203426, f1=0.518348623853211,
best_f1=0.5657015590200446
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.059189632534980774
step: 10, loss: 0.09054331481456757
step: 20, loss: 0.014037393033504486
step: 30, loss: 0.09118708968162537
step: 40, loss: 0.268333375453949
step: 50, loss: 0.3217618763446808
step: 60, loss: 0.03421449288725853
step: 70, loss: 0.04678976163268089
step: 80, loss: 0.030346408486366272
step: 90, loss: 0.1129625141620636
step: 100, loss: 0.14764228463172913
step: 110, loss: 0.018386561423540115
step: 120, loss: 0.05413926765322685
step: 130, loss: 0.07458652555942535
step: 140, loss: 0.03308071568608284
step: 150, loss: 0.05238017067313194
step: 160, loss: 0.11358316987752914
step: 170, loss: 0.06463789194822311
step: 180, loss: 0.21710547804832458
step: 190, loss: 0.04493982717394829
step: 200, loss: 0.0993426963686943
step: 210, loss: 0.03742193803191185
/usr/local/lib/python3.7/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=0.5280898876404494, f1=0.5415019762845851,
best_f1=0.5657015590200446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03195389732718468
step: 10, loss: 0.012023115530610085
step: 20, loss: 0.14388589560985565
step: 30, loss: 0.05393990874290466
step: 40, loss: 0.02454371377825737
step: 50, loss: 0.049394480884075165
step: 60, loss: 0.06843756139278412
step: 70, loss: 0.03189492225646973
step: 80, loss: 0.12861527502536774
step: 90, loss: 0.12306907773017883
step: 100, loss: 0.06603046506643295
step: 110, loss: 0.08964204788208008
step: 120, loss: 0.014200646430253983
step: 130, loss: 0.05927478149533272
step: 140, loss: 0.08804361522197723
```

```
step: 150, loss: 0.05885953828692436
step: 160, loss: 0.1824922263622284
step: 170, loss: 0.011900598183274269
step: 180, loss: 0.006411373615264893
step: 190, loss: 0.09972453117370605
step: 200, loss: 0.03427128866314888
step: 210, loss: 0.18691717088222504
/usr/local/lib/python3.7/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=0.548951048951049, f1=0.5284403669724771,
best_f1=0.5657015590200446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09621010720729828
step: 10, loss: 0.08841152489185333
step: 20, loss: 0.14561910927295685
step: 30, loss: 0.015558785758912563
step: 40, loss: 0.04977957531809807
step: 50, loss: 0.03769414871931076
step: 60, loss: 0.19005385041236877
step: 70, loss: 0.02892349287867546
step: 80, loss: 0.01717534288764
step: 90, loss: 0.05122878775000572
step: 100, loss: 0.08792930841445923
step: 110, loss: 0.07030686736106873
step: 120, loss: 0.132662832736969
step: 130, loss: 0.08105376362800598
step: 140, loss: 0.0033108415082097054
step: 150, loss: 0.034102942794561386
step: 160, loss: 0.06663355231285095
step: 170, loss: 0.0929659977555275
step: 180, loss: 0.004581460263580084
step: 190, loss: 0.03848394751548767
step: 200, loss: 0.01792086847126484
step: 210, loss: 0.08071638643741608
/usr/local/lib/python3.7/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=0.5342465753424659, f1=0.5157699443413729,
best_f1=0.5657015590200446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08826469630002975
step: 10, loss: 0.10699370503425598
step: 20, loss: 0.024570878595113754
```

```
step: 30, loss: 0.04890969395637512
step: 40, loss: 0.18682213127613068
step: 50, loss: 0.03828324377536774
step: 60, loss: 0.2926015257835388
step: 70, loss: 0.02736872062087059
step: 80, loss: 0.020591486245393753
step: 90, loss: 0.057623568922281265
step: 100, loss: 0.010623559355735779
step: 110, loss: 0.00957378838211298
step: 120, loss: 0.005355888046324253
step: 130, loss: 0.020423058420419693
step: 140, loss: 0.00785052590072155
step: 150, loss: 0.009188334457576275
step: 160, loss: 0.13749487698078156
step: 170, loss: 0.04221978038549423
step: 180, loss: 0.0416397750377655
step: 190, loss: 0.12596075236797333
step: 200, loss: 0.07115726917982101
step: 210, loss: 0.04378324747085571
/usr/local/lib/python3.7/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=0.5222929936305732, f1=0.5124716553287981,
best_f1=0.5657015590200446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014938266947865486
step: 10, loss: 0.058909740298986435
step: 20, loss: 0.027099519968032837
step: 30, loss: 0.07541878521442413
step: 40, loss: 0.07068885862827301
step: 50, loss: 0.12935872375965118
step: 60, loss: 0.3383297920227051
step: 70, loss: 0.06727232038974762
step: 80, loss: 0.022799579426646233
step: 90, loss: 0.0819755271077156
step: 100, loss: 0.01953132450580597
step: 110, loss: 0.02233937568962574
step: 120, loss: 0.0075567313469946384
step: 130, loss: 0.1233781874179401
step: 140, loss: 0.02513284981250763
step: 150, loss: 0.012040599249303341
step: 160, loss: 0.001195072429254651
step: 170, loss: 0.12929637730121613
step: 180, loss: 0.01081000454723835
step: 190, loss: 0.041277069598436356
step: 200, loss: 0.008815079927444458
```

step: 210, loss: 0.09215313196182251
/usr/local/lib/python3.7/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=0.5244755244755245, f1=0.5293005671077505,
best_f1=0.5657015590200446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02143867313861847
step: 10, loss: 0.12383875250816345
step: 20, loss: 0.0765041708946228
step: 30, loss: 0.09818779677152634
step: 40, loss: 0.07741951197385788
step: 50, loss: 0.020096732303500175
step: 60, loss: 0.007176543585956097
step: 70, loss: 0.01143505796790123
step: 80, loss: 0.0006289760931394994
step: 90, loss: 0.09508144110441208
step: 100, loss: 0.0013969658175483346
step: 110, loss: 0.011958124116063118
step: 120, loss: 0.012834316119551659
step: 130, loss: 0.0021115345880389214
step: 140, loss: 0.001228840323165059
step: 150, loss: 0.10733999311923981
step: 160, loss: 0.23148535192012787
step: 170, loss: 0.01459378283470869
step: 180, loss: 0.0155977513641119
step: 190, loss: 0.07989058643579483
step: 200, loss: 0.11198893934488297
step: 210, loss: 0.12766021490097046
/usr/local/lib/python3.7/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=0.5322245322245323, f1=0.5054945054945055,
best_f1=0.5657015590200446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012421611696481705
step: 10, loss: 0.045143112540245056
step: 20, loss: 0.0027539250440895557
step: 30, loss: 0.017285946756601334
step: 40, loss: 0.03603852540254593
step: 50, loss: 0.0012021210277453065
step: 60, loss: 0.11545073986053467
step: 70, loss: 0.002287562470883131
step: 80, loss: 0.4468563497066498

```
step: 90, loss: 0.006274957675486803
step: 100, loss: 0.08171573281288147
step: 110, loss: 0.00462091900408268
step: 120, loss: 0.018871551379561424
step: 130, loss: 0.005454206373542547
step: 140, loss: 0.02163657173514366
step: 150, loss: 0.13077887892723083
step: 160, loss: 0.015318487770855427
step: 170, loss: 0.025314418599009514
step: 180, loss: 0.05973192676901817
step: 190, loss: 0.004223293624818325
step: 200, loss: 0.0010765340412035584
step: 210, loss: 0.007183643523603678
/usr/local/lib/python3.7/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=0.5290581162324649, f1=0.5064935064935066,
best_f1=0.5657015590200446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002196087036281824
step: 10, loss: 0.025373412296175957
step: 20, loss: 0.012816214933991432
step: 30, loss: 0.007841403596103191
step: 40, loss: 0.11458437889814377
step: 50, loss: 0.04733128845691681
step: 60, loss: 0.11181391775608063
step: 70, loss: 0.0015637065516784787
step: 80, loss: 0.005233903881162405
step: 90, loss: 0.0014059498207643628
step: 100, loss: 0.008440429344773293
step: 110, loss: 0.006246205419301987
step: 120, loss: 0.19016867876052856
step: 130, loss: 0.021424056962132454
step: 140, loss: 0.005227480083703995
step: 150, loss: 0.004572426900267601
step: 160, loss: 0.1286865472793579
step: 170, loss: 0.06122981384396553
step: 180, loss: 0.002994810463860631
step: 190, loss: 0.033648282289505005
step: 200, loss: 0.010066860355436802
step: 210, loss: 0.02260311134159565
/usr/local/lib/python3.7/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=0.49350649350649345, f1=0.4954545454545455,
best_f1=0.5657015590200446
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015324595151469111
step: 10, loss: 0.008349242620170116
step: 20, loss: 0.031061185523867607
step: 30, loss: 0.035978902131319046
step: 40, loss: 0.006416153162717819
step: 50, loss: 0.003222191007807851
step: 60, loss: 0.009537297300994396
step: 70, loss: 0.04846707358956337
step: 80, loss: 0.008993488736450672
step: 90, loss: 0.008395128883421421
step: 100, loss: 0.0285254567861557
step: 110, loss: 0.004690655041486025
step: 120, loss: 0.007716774009168148
step: 130, loss: 0.0036863330751657486
step: 140, loss: 0.002851405180990696
step: 150, loss: 0.0006738496595062315
step: 160, loss: 0.004184602294117212
step: 170, loss: 0.001715374644845724
step: 180, loss: 0.1377231627702713
step: 190, loss: 0.0030361160170286894
step: 200, loss: 0.011817210353910923
step: 210, loss: 0.01601319946348667
/usr/local/lib/python3.7/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=0.5257731958762887, f1=0.5243445692883895,
best_f1=0.5657015590200446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0029894239269196987
step: 10, loss: 0.002819381421431899
step: 20, loss: 0.015630895271897316
step: 30, loss: 0.004016614984720945
step: 40, loss: 0.0033020381815731525
step: 50, loss: 0.002130730077624321
step: 60, loss: 0.07325707376003265
step: 70, loss: 0.12207530438899994
step: 80, loss: 0.09146357327699661
step: 90, loss: 0.09142038226127625
step: 100, loss: 0.009026609361171722
step: 110, loss: 0.006602159701287746
step: 120, loss: 0.0019273092038929462
step: 130, loss: 0.0012471299851313233
step: 140, loss: 0.014974481426179409
```

```
step: 150, loss: 0.059120893478393555
step: 160, loss: 0.043439626693725586
step: 170, loss: 0.03018122725188732
step: 180, loss: 0.0024252147413790226
step: 190, loss: 0.00901097059249878
step: 200, loss: 0.003287229686975479
step: 210, loss: 0.01529285591095686
/usr/local/lib/python3.7/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=0.5309090909090909, f1=0.5188118811881188,
best_f1=0.5657015590200446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003575169248506427
step: 10, loss: 0.009021961130201817
step: 20, loss: 0.0032268965151160955
step: 30, loss: 0.001462298328988254
step: 40, loss: 0.0014923736453056335
step: 50, loss: 0.006323367357254028
step: 60, loss: 0.036937206983566284
step: 70, loss: 0.0033922891598194838
step: 80, loss: 0.0009234707104042172
step: 90, loss: 0.007280552759766579
step: 100, loss: 0.0007418326567858458
step: 110, loss: 0.007555793970823288
step: 120, loss: 0.08764563500881195
step: 130, loss: 0.002370232716202736
step: 140, loss: 0.0037569350097328424
step: 150, loss: 0.006188997998833656
step: 160, loss: 0.014014048501849174
step: 170, loss: 0.04706953838467598
step: 180, loss: 0.0007820118335075676
step: 190, loss: 0.006450004875659943
step: 200, loss: 0.004432190675288439
step: 210, loss: 0.06126760318875313
/usr/local/lib/python3.7/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=0.5245901639344263, f1=0.505643340857788,
best_f1=0.5657015590200446
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.")
2293it [00:06, 341.18it/s]
load_f1 = 0.5595238095238095
real_f1 = 0.5593561368209256
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, 258.02it/s]
```

1.1.5 Walmart-Amazon - Running the matcher

```python
# 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 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: 0.8646589517593384
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.16717839241027832
step: 20, loss: 0.1551298350095749
step: 30, loss: 0.49696946144104004
step: 40, loss: 0.24942858517169952
step: 50, loss: 0.3039184510707855
step: 60, loss: 0.35644033551216125
step: 70, loss: 0.17397822439670563
step: 80, loss: 0.5061125159263611
step: 90, loss: 0.23906217515468597
step: 100, loss: 0.2191791981458664
step: 110, loss: 0.23251672089099884
step: 120, loss: 0.4100545346736908
step: 130, loss: 0.34607771039009094
step: 140, loss: 0.3027927279472351
step: 150, loss: 0.2896427810192108
step: 160, loss: 0.24292732775211334
step: 170, loss: 0.42000386118888855
step: 180, loss: 0.27107667922973633
step: 190, loss: 0.122832752764225
/usr/local/lib/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=0.5427350427350427, f1=0.5828092243186582,
best_f1=0.5828092243186582
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26785019040107727
step: 10, loss: 0.03249606117606163
step: 20, loss: 0.06931190937757492
step: 30, loss: 0.18216393887996674
step: 40, loss: 0.5184376835823059
step: 50, loss: 0.32975640892982483
step: 60, loss: 0.2168717086315155
```

```
step: 70, loss: 0.20494131743907928
step: 80, loss: 0.12069098651409149
step: 90, loss: 0.1422763168811798
step: 100, loss: 0.1702488511800766
step: 110, loss: 0.11922195553779602
step: 120, loss: 0.16003236174583435
step: 130, loss: 0.12622323632240295
step: 140, loss: 0.2774313688278198
step: 150, loss: 0.13679012656211853
step: 160, loss: 0.02615688182413578
step: 170, loss: 0.15606176853179932
step: 180, loss: 0.07951439917087555
step: 190, loss: 0.11524757742881775
/usr/local/lib/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=0.7365591397849462, f1=0.7472527472527473,
best_f1=0.7472527472527473
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10572968423366547
step: 10, loss: 0.15128503739833832
step: 20, loss: 0.1013847291469574
step: 30, loss: 0.030344601720571518
step: 40, loss: 0.03516559302806854
step: 50, loss: 0.1832873523235321
step: 60, loss: 0.09453020244836807
step: 70, loss: 0.05907725542783737
step: 80, loss: 0.19521751999855042
step: 90, loss: 0.04469640925526619
step: 100, loss: 0.1458549052476883
step: 110, loss: 0.08945414423942566
step: 120, loss: 0.010397039353847504
step: 130, loss: 0.10626243054866791
step: 140, loss: 0.06258173286914825
step: 150, loss: 0.13156820833683014
step: 160, loss: 0.06556685268878937
step: 170, loss: 0.038758475333452225
step: 180, loss: 0.07814978063106537
step: 190, loss: 0.140813410282135
/usr/local/lib/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=0.7586206896551724, f1=0.7814207650273224,
best_f1=0.7814207650273224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02231748029589653
step: 10, loss: 0.07000701129436493
step: 20, loss: 0.021134084090590477
step: 30, loss: 0.0638008862733841
step: 40, loss: 0.004177554976195097
step: 50, loss: 0.030120549723505974
step: 60, loss: 0.099740169942379
step: 70, loss: 0.006972557865083218
step: 80, loss: 0.025226781144738197
step: 90, loss: 0.09242164343595505
step: 100, loss: 0.01024064514786005
step: 110, loss: 0.11957104504108429
step: 120, loss: 0.04775944724678993
step: 130, loss: 0.185654416680336
step: 140, loss: 0.08703280240297318
step: 150, loss: 0.015983564779162407
step: 160, loss: 0.06917352229356766
step: 170, loss: 0.06487391889095306
step: 180, loss: 0.07218684256076813
step: 190, loss: 0.02871982753276825
/usr/local/lib/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=0.7647058823529412, f1=0.7507002801120448,
best_f1=0.7507002801120448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0519254244863987
step: 10, loss: 0.04953256994485855
step: 20, loss: 0.09958572685718536
step: 30, loss: 0.03174501657485962
step: 40, loss: 0.0734737366437912
step: 50, loss: 0.02425030805170536
step: 60, loss: 0.027677522972226143
```

```
step: 70, loss: 0.011140482500195503
step: 80, loss: 0.020276842638850212
step: 90, loss: 0.0029601554851979017
step: 100, loss: 0.009785635396838188
step: 110, loss: 0.0012237756745889783
step: 120, loss: 0.003601408563554287
step: 130, loss: 0.07230661809444427
step: 140, loss: 0.00579103734344244
step: 150, loss: 0.048798155039548874
step: 160, loss: 0.01927792653441429
step: 170, loss: 0.027499467134475708
step: 180, loss: 0.10039161145687103
step: 190, loss: 0.0632133036851883
/usr/local/lib/python3.7/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=0.76, f1=0.7544910179640719, best_f1=0.7507002801120448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014866724610328674
step: 10, loss: 0.00951592717319727
step: 20, loss: 0.008302191272377968
step: 30, loss: 0.002949533285573125
step: 40, loss: 0.033779554069042206
step: 50, loss: 0.021331248804926872
step: 60, loss: 0.02388218604028225
step: 70, loss: 0.11049901694059372
step: 80, loss: 0.13114772737026215
step: 90, loss: 0.10747137665748596
step: 100, loss: 0.011755642481148243
step: 110, loss: 0.015653571113944054
step: 120, loss: 0.010810093954205513
step: 130, loss: 0.015424714423716068
step: 140, loss: 0.00311599625274539
step: 150, loss: 0.058317724615335464
step: 160, loss: 0.00964093953371048
step: 170, loss: 0.04330061748623848
step: 180, loss: 0.021823396906256676
step: 190, loss: 0.006839402485638857
/usr/local/lib/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=0.7715736040609137, f1=0.7473684210526317,
```

```
best_f1=0.7473684210526317
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0037314551882445812
step: 10, loss: 0.06980336457490921
step: 20, loss: 0.010927959345281124
step: 30, loss: 0.010638738051056862
step: 40, loss: 0.01829611137509346
step: 50, loss: 0.051428597420454025
step: 60, loss: 0.03568826615810394
step: 70, loss: 0.0012353723868727684
step: 80, loss: 0.03570324555039406
step: 90, loss: 0.0034099200274795294
step: 100, loss: 0.007636707276105881
step: 110, loss: 0.009933743625879288
step: 120, loss: 0.004765671212226152
step: 130, loss: 0.014872090891003609
step: 140, loss: 0.0026068687438964844
step: 150, loss: 0.023804349824786186
step: 160, loss: 0.0018027009209617972
step: 170, loss: 0.009533317759633064
step: 180, loss: 0.07416708022356033
step: 190, loss: 0.01977355219423771
/usr/local/lib/python3.7/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=0.7412935323383084, f1=0.7171717171717172,
best_f1=0.7473684210526317
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002200792310759425
step: 10, loss: 0.002010094001889229
step: 20, loss: 0.0007862868951633573
step: 30, loss: 0.0032510256860405207
step: 40, loss: 0.01758253388106823
step: 50, loss: 0.0023677425924688578
step: 60, loss: 0.015347070060670376
step: 70, loss: 0.014469582587480545
step: 80, loss: 0.12291179597377777
step: 90, loss: 0.01156620867550373
step: 100, loss: 0.005645877216011286
step: 110, loss: 0.0031831893138587475
step: 120, loss: 0.07297417521476746
step: 130, loss: 0.013414415530860424
step: 140, loss: 0.00411256542429328
step: 150, loss: 0.0017197455745190382
```

```
step: 160, loss: 0.043844904750585556
step: 170, loss: 0.0010702430736273527
step: 180, loss: 0.017779570072889328
step: 190, loss: 0.042854927480220795
/usr/local/lib/python3.7/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=0.7639257294429709, f1=0.7252747252747253,
best_f1=0.7473684210526317
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003724508685991168
step: 10, loss: 0.0011601010337471962
step: 20, loss: 0.021465810015797615
step: 30, loss: 0.08991484344005585
step: 40, loss: 0.012788872234523296
step: 50, loss: 0.0010957899503409863
step: 60, loss: 0.0007254349184222519
step: 70, loss: 0.02083154395222664
step: 80, loss: 0.0012702231761068106
step: 90, loss: 0.005447313189506531
step: 100, loss: 0.004174943547695875
step: 110, loss: 0.034552205353975296
step: 120, loss: 0.22908703982830048
step: 130, loss: 0.007212443742901087
step: 140, loss: 0.002022639848291874
step: 150, loss: 0.018292739987373352
step: 160, loss: 0.0049524144269526005
step: 170, loss: 0.0007796087302267551
step: 180, loss: 0.011165500618517399
step: 190, loss: 0.04175795242190361
/usr/local/lib/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=0.7760416666666666, f1=0.7465940054495912,
best_f1=0.7465940054495912
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0451299250125885
step: 10, loss: 0.007745302747935057
step: 20, loss: 0.00046042705071158707
step: 30, loss: 0.0011182614834979177
```

```
step: 40, loss: 0.0005009905435144901
step: 50, loss: 0.00036316245677880943
step: 60, loss: 0.0020797522738575935
step: 70, loss: 0.0023233087266772666
step: 80, loss: 0.004467529710382223
step: 90, loss: 0.001763631938956678
step: 100, loss: 0.009219043888151646
step: 110, loss: 0.02827899344265461
step: 120, loss: 0.00848125759512186
step: 130, loss: 0.01687423512339592
step: 140, loss: 0.0009385923622176051
step: 150, loss: 0.0008048051968216896
step: 160, loss: 0.00915268063545227
step: 170, loss: 0.003382124938070774
step: 180, loss: 0.0793214812874794
step: 190, loss: 0.016034893691539764
/usr/local/lib/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=0.7877094972067039, f1=0.7420289855072463,
best_f1=0.7420289855072463
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011209151707589626
step: 10, loss: 0.001509023131802678
step: 20, loss: 0.21307533979415894
step: 30, loss: 0.0113410186022252007
step: 40, loss: 0.0007257999968715012
step: 50, loss: 0.004558270797133446
step: 60, loss: 0.001762081403285265
step: 70, loss: 0.0018867457984015346
step: 80, loss: 0.0072796703316271305
step: 90, loss: 0.08447688072919846
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 100, loss: 0.0004903247463516891
step: 110, loss: 0.0006197018665261567
step: 120, loss: 0.010694054886698723
step: 130, loss: 0.004854253027588129
step: 140, loss: 0.001039988943375647
step: 150, loss: 0.0019580263178795576
step: 160, loss: 0.002039248589426279
step: 170, loss: 0.0027724909596145153
step: 180, loss: 0.05005193129181862
```

step: 190, loss: 0.01988079771399498
/usr/local/lib/python3.7/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=0.7862796833773087, f1=0.7747252747252747,
best_f1=0.7420289855072463
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014582639560103416
step: 10, loss: 0.09650695323944092
step: 20, loss: 0.04511464387178421
step: 30, loss: 0.014014611020684242
step: 40, loss: 0.016879037022590637
step: 50, loss: 0.03714675456285477
step: 60, loss: 0.001046335557475686
step: 70, loss: 0.0009095954010263085
step: 80, loss: 0.0013531692093238235
step: 90, loss: 0.02411174401640892
step: 100, loss: 0.05789029970765114
step: 110, loss: 0.0005575526738539338
step: 120, loss: 0.0010651114862412214
step: 130, loss: 0.0007383501506410539
step: 140, loss: 0.0012414483353495598
step: 150, loss: 0.0026199237909168005
step: 160, loss: 0.0017928924644365907
step: 170, loss: 0.019903911277651787
step: 180, loss: 0.0023898924700915813
step: 190, loss: 0.036890909075737
/usr/local/lib/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=0.7893333333333333, f1=0.7450980392156863,
best_f1=0.7450980392156863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005249704699963331
step: 10, loss: 0.002787957200780511
step: 20, loss: 0.09188951551914215
step: 30, loss: 0.001492544892244041
step: 40, loss: 0.0007289904169738293
step: 50, loss: 0.0005708811804652214
step: 60, loss: 0.00132374360691756

```
step: 70, loss: 0.0048091295175254345
step: 80, loss: 0.0005690394318662584
step: 90, loss: 0.0016674912767484784
step: 100, loss: 0.000410582113545388
step: 110, loss: 0.021487189456820488
step: 120, loss: 0.0017223614268004894
step: 130, loss: 0.020536210387945175
step: 140, loss: 0.004229600075632334
step: 150, loss: 0.0050984499976603893
step: 160, loss: 0.003517101751640439
step: 170, loss: 0.0016898349858820438
step: 180, loss: 0.002291338052600622
step: 190, loss: 0.004176876042038202
/usr/local/lib/python3.7/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=0.768421052631579, f1=0.7317073170731707,
best_f1=0.7450980392156863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006389811635017395
step: 10, loss: 0.0010906660463660955
step: 20, loss: 0.0012712378520518541
step: 30, loss: 0.0011804995592683554
step: 40, loss: 0.0006688644061796367
step: 50, loss: 0.0010095880134031177
step: 60, loss: 0.09546145796775818
step: 70, loss: 0.16253244876861572
step: 80, loss: 0.0015500427689403296
step: 90, loss: 0.001537919044494629
step: 100, loss: 0.00115659786388278
step: 110, loss: 0.01792813278734684
step: 120, loss: 0.0016085469396784902
step: 130, loss: 0.0020391521975398064
step: 140, loss: 0.00040969831752590835
step: 150, loss: 0.0005356392357498407
step: 160, loss: 0.00043940715841017663
step: 170, loss: 0.004203427117317915
step: 180, loss: 0.003024816047400236
step: 190, loss: 0.0003591330023482442
/usr/local/lib/python3.7/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=0.7716535433070866, f1=0.7520435967302452,
best_f1=0.7450980392156863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004663774743676186
step: 10, loss: 0.0006853415980003774
step: 20, loss: 0.0010402984917163849
step: 30, loss: 0.0007423001225106418
step: 40, loss: 0.002772783860564232
step: 50, loss: 0.003759292419999838
step: 60, loss: 0.00431208498775959
step: 70, loss: 0.0010014149593189359
step: 80, loss: 0.0009414759115315974
step: 90, loss: 0.0024368558079004288
step: 100, loss: 0.0031979898922145367
step: 110, loss: 0.0014003493124619126
step: 120, loss: 0.0008723728242330253
step: 130, loss: 0.0015621253987774253
step: 140, loss: 0.002157042734324932
step: 150, loss: 0.0037085795775055885
step: 160, loss: 0.002783567411825061
step: 170, loss: 0.0024072567466646433
step: 180, loss: 0.0010783903999254107
step: 190, loss: 0.0021055834367871284
/usr/local/lib/python3.7/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=0.7676240208877285, f1=0.7419354838709677,
best_f1=0.7450980392156863
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.")
2049it [00:09, 223.41it/s]
load_f1 = 0.7382198952879582
real_f1 = 0.7189873417721518
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, 257.58it/s]
```

1.2 DA TEXTUAL

1.2.1 Abt-Buy - Running the matcher

```python
# 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 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: 0.8526486754417419
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.23128478229045868
step: 20, loss: 0.15582528710365295
step: 30, loss: 0.23072081804275513
step: 40, loss: 0.3106083869934082

```

```
step: 50, loss: 0.3780853748321533
step: 60, loss: 0.4591792821884155
step: 70, loss: 0.3039804995059967
step: 80, loss: 0.2517528533935547
step: 90, loss: 0.3900551199913025
step: 100, loss: 0.23235853016376495
step: 110, loss: 0.17866148054599762
step: 120, loss: 0.5267549753189087
step: 130, loss: 0.4011763334274292
step: 140, loss: 0.37492242455482483
step: 150, loss: 0.15663744509220123
step: 160, loss: 0.17291457951068878
step: 170, loss: 0.0719032809138298
/usr/local/lib/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=0.6870588235294117, f1=0.6574074074074074,
best_f1=0.6574074074074074
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2786271572113037
step: 10, loss: 0.12699012458324432
step: 20, loss: 0.222061425447464
step: 30, loss: 0.07099416851997375
step: 40, loss: 0.12444952130317688
step: 50, loss: 0.16994339227676392
step: 60, loss: 0.06668481975793839
step: 70, loss: 0.1628076285123825
step: 80, loss: 0.11516132950782776
step: 90, loss: 0.1667134314775467
step: 100, loss: 0.1528300940990448
step: 110, loss: 0.2370034009218216
step: 120, loss: 0.02913692593574524
step: 130, loss: 0.08103219419717789
step: 140, loss: 0.08269500732421875
step: 150, loss: 0.17303681373596191
step: 160, loss: 0.13748231530189514
step: 170, loss: 0.2944530248641968
/usr/local/lib/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=0.7505827505827506, f1=0.7455357142857144,
best_f1=0.7455357142857144
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03323623165488243
step: 10, loss: 0.03677089884877205
step: 20, loss: 0.07307126373052597
step: 30, loss: 0.1295015811920166
step: 40, loss: 0.023311790078878403
step: 50, loss: 0.09214877337217331
step: 60, loss: 0.23942379653453827
step: 70, loss: 0.05992730334401131
step: 80, loss: 0.1796034872531891
step: 90, loss: 0.07007189840078354
step: 100, loss: 0.02220047451555729
step: 110, loss: 0.031597308814525604
step: 120, loss: 0.01218139473348856
step: 130, loss: 0.09500839561223984
step: 140, loss: 0.009513180702924728
step: 150, loss: 0.12631846964359283
step: 160, loss: 0.04106902331113815
step: 170, loss: 0.09410762786865234
/usr/local/lib/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=0.754880694143167, f1=0.7372881355932204,
best_f1=0.7372881355932204
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.034369293600320816
step: 10, loss: 0.2345159351825714
step: 20, loss: 0.0862630233168602
step: 30, loss: 0.04323330149054527
step: 40, loss: 0.0491500161588192
step: 50, loss: 0.0034970720298588276
step: 60, loss: 0.06316320598125458
step: 70, loss: 0.01254928857088089
step: 80, loss: 0.013026638887822628
step: 90, loss: 0.014226026833057404
step: 100, loss: 0.008632024750113487

```
step: 110, loss: 0.026734767481684685
step: 120, loss: 0.1694415807723999
step: 130, loss: 0.04333160072565079
step: 140, loss: 0.022968821227550507
step: 150, loss: 0.002096431329846382
step: 160, loss: 0.03869787976145744
step: 170, loss: 0.021206118166446686
/usr/local/lib/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=0.7692307692307693, f1=0.7373068432671083,
best_f1=0.7373068432671083
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11541008949279785
step: 10, loss: 0.15501460433006287
step: 20, loss: 0.01357860304415226
step: 30, loss: 0.037789374589920044
step: 40, loss: 0.02015201561152935
step: 50, loss: 0.043204374611377716
step: 60, loss: 0.010879195295274258
step: 70, loss: 0.01743628829717636
step: 80, loss: 0.03757958486676216
step: 90, loss: 0.02048325352370739
step: 100, loss: 0.04242134839296341
step: 110, loss: 0.17512120306491852
step: 120, loss: 0.05896424874663353
step: 130, loss: 0.02165118418633938
step: 140, loss: 0.03886161744594574
step: 150, loss: 0.002854754449799657
step: 160, loss: 0.005712655372917652
step: 170, loss: 0.05851978436112404
/usr/local/lib/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=0.7738927738927739, f1=0.7439824945295405,
best_f1=0.7439824945295405
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022719070315361023
step: 10, loss: 0.02178596332669258
step: 20, loss: 0.024426670745015144
step: 30, loss: 0.0752454400062561
step: 40, loss: 0.03564170002937317
step: 50, loss: 0.013114352710545063
step: 60, loss: 0.02680104970932007
step: 70, loss: 0.011580538004636765
step: 80, loss: 0.032132551074028015
step: 90, loss: 0.04634160175919533
step: 100, loss: 0.005770395044237375
step: 110, loss: 0.06272904574871063
step: 120, loss: 0.032096587121486664
step: 130, loss: 0.19241511821746826
step: 140, loss: 0.016561036929488182
step: 150, loss: 0.18928217887878418
step: 160, loss: 0.03548189997673035
step: 170, loss: 0.002285139635205269
/usr/local/lib/python3.7/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=0.7314578005115089, f1=0.7855421686746988,
best_f1=0.7439824945295405
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003312360728159547
step: 10, loss: 0.0010322119342163205
step: 20, loss: 0.004786084406077862
step: 30, loss: 0.010052423924207687
step: 40, loss: 0.011035718955099583
step: 50, loss: 0.003940480295568705
step: 60, loss: 0.2407780885696411
step: 70, loss: 0.01255244854837656
step: 80, loss: 0.007479558698832989
step: 90, loss: 0.012263420969247818
step: 100, loss: 0.015228752046823502
step: 110, loss: 0.00028096564346924424
step: 120, loss: 0.26812735199928284
step: 130, loss: 0.09961015731096268
step: 140, loss: 0.015852520242333412
step: 150, loss: 0.05447008088231087
step: 160, loss: 0.09447143226861954
step: 170, loss: 0.054696694016456604
/usr/local/lib/python3.7/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=0.7536945812807883, f1=0.7499999999999999,
best_f1=0.7439824945295405
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07959015667438507
step: 10, loss: 0.025615105405449867
step: 20, loss: 0.004972720053046942
step: 30, loss: 0.01088624820113182
step: 40, loss: 0.10034680366516113
step: 50, loss: 0.0026740562170743942
step: 60, loss: 0.03963148593902588
step: 70, loss: 0.024306105449795723
step: 80, loss: 0.012733208946883678
step: 90, loss: 0.010442838072776794
step: 100, loss: 0.013782468624413013
step: 110, loss: 0.049777694046497345
step: 120, loss: 0.0028296697419136763
step: 130, loss: 0.06467566639184952
step: 140, loss: 0.004903649911284447
step: 150, loss: 0.019235312938690186
step: 160, loss: 0.010738805867731571
step: 170, loss: 0.003455876372754574
/usr/local/lib/python3.7/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=0.7595505617977529, f1=0.7308533916849015,
best_f1=0.7439824945295405
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000424739409936592
step: 10, loss: 0.07383007556200027
step: 20, loss: 0.004016650840640068
step: 30, loss: 0.06338387727737427
step: 40, loss: 0.08859237283468246
step: 50, loss: 0.010529603809118271
step: 60, loss: 0.002102839294821024
step: 70, loss: 0.035731855779886246
step: 80, loss: 0.022619346156716347
step: 90, loss: 0.004638251382857561
step: 100, loss: 0.017615433782339096
step: 110, loss: 0.006707628723233938
step: 120, loss: 0.00515371048822999
step: 130, loss: 0.002402913523837924
step: 140, loss: 0.0003530229441821575
step: 150, loss: 0.08967716246843338
step: 160, loss: 0.019909823313355446
```

```
step: 170, loss: 0.014371838420629501
/usr/local/lib/python3.7/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=0.7427055702917773, f1=0.7661691542288557,
best_f1=0.7439824945295405
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005755330203101039
step: 10, loss: 0.013965575955808163
step: 20, loss: 0.0018768529407680035
step: 30, loss: 0.007522656116634607
step: 40, loss: 0.03281555324792862
step: 50, loss: 0.004177193623036146
step: 60, loss: 0.00164191541261971
step: 70, loss: 0.004078722558915615
step: 80, loss: 0.0006528893136419356
step: 90, loss: 0.0608462318778038
step: 100, loss: 0.01977379433810711
step: 110, loss: 0.0007845698855817318
step: 120, loss: 0.002913457341492176
step: 130, loss: 0.008469638414680958
step: 140, loss: 0.0555715337395668
step: 150, loss: 0.04332157224416733
step: 160, loss: 0.016538886353373528
step: 170, loss: 0.0009257450583390892
/usr/local/lib/python3.7/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=0.7347931873479318, f1=0.7348837209302326,
best_f1=0.7439824945295405
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0025663580745458603
step: 10, loss: 0.03377115726470947
step: 20, loss: 0.11246834695339203
step: 30, loss: 0.002253541722893715
step: 40, loss: 0.012145120650529861
step: 50, loss: 0.0007131862221285701
step: 60, loss: 0.004188460763543844
step: 70, loss: 0.01535727083683014
step: 80, loss: 0.0012358683161437511
step: 90, loss: 0.0032564697321504354
step: 100, loss: 0.0007837974117137492
step: 110, loss: 0.0008297436288557947
step: 120, loss: 0.003470313036814332
```

```
step: 130, loss: 0.018550682812929153
step: 140, loss: 0.03728891536593437
step: 150, loss: 0.07077355682849884
step: 160, loss: 0.0018593244021758437
step: 170, loss: 0.0720750167965889
/usr/local/lib/python3.7/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=0.7461139896373058, f1=0.7673267326732675,
best_f1=0.7439824945295405
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0032610404305160046
step: 10, loss: 0.0064725130796432495
step: 20, loss: 0.013465550728142262
step: 30, loss: 0.0013436625013127923
step: 40, loss: 0.00024954858236014843
step: 50, loss: 0.0007989084115251899
step: 60, loss: 0.0005040210671722889
step: 70, loss: 0.03034796193242073
step: 80, loss: 0.019029734656214714
step: 90, loss: 0.0009094459237530828
step: 100, loss: 0.006573397666215897
step: 110, loss: 0.0006195370224304497
step: 120, loss: 0.06373513489961624
step: 130, loss: 0.001208007801324129
step: 140, loss: 0.0002806837728712708
step: 150, loss: 0.05365949869155884
step: 160, loss: 0.01634848304092884
step: 170, loss: 0.0003284290141891688
/usr/local/lib/python3.7/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=0.7382198952879581, f1=0.7780548628428928,
best_f1=0.7439824945295405
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0059138089418411255
step: 10, loss: 0.09817575663328171
step: 20, loss: 0.008231832645833492
step: 30, loss: 0.08504170924425125
step: 40, loss: 0.00011406542762415484
step: 50, loss: 0.04369513317942619
step: 60, loss: 0.00036910182097926736
step: 70, loss: 0.013582224026322365
step: 80, loss: 0.0021465087775141
```

step: 90, loss: 0.0006633788580074906
step: 100, loss: 8.196472481358796e-05
step: 110, loss: 0.0010090277064591646
step: 120, loss: 0.0005646725767292082
step: 130, loss: 0.010769396089017391
step: 140, loss: 0.00023071911709848791
step: 150, loss: 0.0003974501451011747
step: 160, loss: 0.05770888179540634
step: 170, loss: 0.0011225015623494983
/usr/local/lib/python3.7/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=0.7477064220183486, f1=0.7488986784140969,
best_f1=0.7439824945295405
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002051994437351823
step: 10, loss: 0.00014624622417613864
step: 20, loss: 0.0025307838805019855
step: 30, loss: 0.0003559321630746126
step: 40, loss: 0.0012185215018689632
step: 50, loss: 0.0008685786160640419
step: 60, loss: 0.00023766023514326662
step: 70, loss: 0.0012317111250013113
step: 80, loss: 0.00037461568717844784
step: 90, loss: 0.0007699029520153999
step: 100, loss: 0.000672042602673173
step: 110, loss: 0.0003761483239941299
step: 120, loss: 0.001170033123344183
step: 130, loss: 0.001777534605935216
step: 140, loss: 0.00026672281092032790
step: 150, loss: 0.0011187340132892132
step: 160, loss: 0.02525780163705349
step: 170, loss: 0.00015701384108979255
/usr/local/lib/python3.7/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=0.7437185929648242, f1=0.7751196172248804,
best_f1=0.7439824945295405
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005864868289791048
step: 10, loss: 0.0783575102686882
step: 20, loss: 0.003675221698358655
step: 30, loss: 0.004234889056533575
step: 40, loss: 0.0010284106247127056

```
step: 50, loss: 0.00243912055157125
step: 60, loss: 8.915736543713138e-05
step: 70, loss: 0.00026927253929898143
step: 80, loss: 0.001472362200729549
step: 90, loss: 0.00042166817001998425
step: 100, loss: 0.0005870566819794476
step: 110, loss: 0.0005548496847040951
step: 120, loss: 0.002183922566473484
step: 130, loss: 0.010450503788888454
step: 140, loss: 0.00012528910883702338
step: 150, loss: 0.014112260192632675
step: 160, loss: 0.0001060732247424312
step: 170, loss: 0.0006039640284143388
```
/usr/local/lib/python3.7/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=0.7392405063291139, f1=0.7759036144578313,
best_f1=0.7439824945295405
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.")

                                    99
```

```
1916it [00:05, 327.88it/s]
load_f1 = 0.46096654275092935
real_f1 = 0.4205298013245033
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, 254.83it/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, 489kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 1.22MB/s]
Downloading: 100% 268M/268M [00:13<00:00, 20.2MB/s]
```

```
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: 0.8214190602302551
step: 10, loss: 0.4678594470024109
step: 20, loss: 0.6117331981658936
step: 30, loss: 0.47506770491600037
step: 40, loss: 0.26385822892189026
step: 50, loss: 0.0984354317188263
step: 60, loss: 0.19238387048244476
step: 70, loss: 0.1620909422636032
step: 80, loss: 0.25812894105911255
step: 90, loss: 0.10826946049928665
step: 100, loss: 0.12439142912626266
step: 110, loss: 0.13836432993412018
```

```
step: 120, loss: 0.057610560208559036
step: 130, loss: 0.008220772258937359
step: 140, loss: 0.059066906571388245
step: 150, loss: 0.16268980503082275
step: 160, loss: 0.3064647316932678
step: 170, loss: 0.013817987404763699
step: 180, loss: 0.009215309284627438
step: 190, loss: 0.13238182663917542
step: 200, loss: 0.005334602668881416
step: 210, loss: 0.013616163283586502
step: 220, loss: 0.03790375590324402
step: 230, loss: 0.023489899933338165
/usr/local/lib/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=0.9503858875413451, f1=0.9611542730299667,
best_f1=0.9611542730299667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08029770106077194
step: 10, loss: 0.09342806786298752
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.08736647665500641
step: 30, loss: 0.021192610263824463
step: 40, loss: 0.06700550764799118
step: 50, loss: 0.00868481770157814
step: 60, loss: 0.01643669232726097
step: 70, loss: 0.12672601640224457
step: 80, loss: 0.017964530736207962
step: 90, loss: 0.013343920931220055
step: 100, loss: 0.12912990152835846
step: 110, loss: 0.023595919832587242
step: 120, loss: 0.010442540980875492
step: 130, loss: 0.0020908562000840902
step: 140, loss: 0.03725498914718628
step: 150, loss: 0.017477715387940407
step: 160, loss: 0.041180532425642014
step: 170, loss: 0.09213203191757202
step: 180, loss: 0.027288159355521202
step: 190, loss: 0.08832257986068726
step: 200, loss: 0.14108073711395264
step: 210, loss: 0.05621061474084854
step: 220, loss: 0.003575008362531662
```

```
step: 230, loss: 0.010048380121588707
/usr/local/lib/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=0.9670828603859251, f1=0.9636363636363636,
best_f1=0.9636363636363636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04810977727174759
step: 10, loss: 0.12439858913421631
step: 20, loss: 0.014126980677247047
step: 30, loss: 0.028710445389151573
step: 40, loss: 0.018566714599728584
step: 50, loss: 0.01303945854306221
step: 60, loss: 0.1377752125263214
step: 70, loss: 0.0017799811903387308
step: 80, loss: 0.0041544693522155285
step: 90, loss: 0.0520608127117157
step: 100, loss: 0.03806917369365692
step: 110, loss: 0.027839532122015953
step: 120, loss: 0.04464539885520935
step: 130, loss: 0.002161365933716297
step: 140, loss: 0.003651185194030404
step: 150, loss: 0.003742339089512825
step: 160, loss: 0.0013020781334489584
step: 170, loss: 0.013573052361607552
step: 180, loss: 0.0026695129927247763
step: 190, loss: 0.04483631253242493
step: 200, loss: 0.009718026965856552
step: 210, loss: 0.002511250087991357
step: 220, loss: 0.006056902930140495
step: 230, loss: 0.10551323741674423
/usr/local/lib/python3.7/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=0.9627959413754228, f1=0.9728506787330317,
best_f1=0.9636363636363636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008304571732878685
step: 10, loss: 0.008977465331554413
step: 20, loss: 0.0035879239439964294
```

```
step: 30, loss: 0.002634296426549554
step: 40, loss: 0.0034169543068856
step: 50, loss: 0.0015027981717139482
step: 60, loss: 0.0005563963786698878
step: 70, loss: 0.0026287275832146406
step: 80, loss: 0.12929312884807587
step: 90, loss: 0.0012075876584276557
step: 100, loss: 0.008089362643659115
step: 110, loss: 0.011974669061601162
step: 120, loss: 0.13710948824882507
step: 130, loss: 0.046262260526418686
step: 140, loss: 0.0027443766593933105
step: 150, loss: 0.005389368627220392
step: 160, loss: 0.0037178697530180216
step: 170, loss: 0.0018665831303223968
step: 180, loss: 0.13673292100429535
step: 190, loss: 0.0011288875248283148
step: 200, loss: 0.001992024714127183
step: 210, loss: 0.00310060506314039
step: 220, loss: 0.005946421530097723
step: 230, loss: 0.000590079347603023
/usr/local/lib/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=0.9686800894854586, f1=0.9697648376259798,
best_f1=0.9697648376259798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010201571276411414
step: 10, loss: 0.0007962583913467824
step: 20, loss: 0.0008962334832176566
step: 30, loss: 0.0006065806373953819
step: 40, loss: 0.0011577893747016788
step: 50, loss: 0.0005063908756710589
step: 60, loss: 0.0005793105228804052
step: 70, loss: 0.0002446560247335583
step: 80, loss: 0.007674197666347027
step: 90, loss: 0.0004217060632072389
step: 100, loss: 0.006588185206055641
step: 110, loss: 0.0020635914988815784
step: 120, loss: 0.06549163162708282
step: 130, loss: 0.01495447289198637
step: 140, loss: 0.0007552210008725524
```

```
step: 150, loss: 0.00036914952215738595
step: 160, loss: 0.10839713364839554
step: 170, loss: 0.03469376266002655
step: 180, loss: 0.0036687764804810286
step: 190, loss: 0.0009001228027045727
step: 200, loss: 0.0037207696586847305
step: 210, loss: 0.001884637400507927
step: 220, loss: 0.0034631460439413786
step: 230, loss: 0.009877449832856655
/usr/local/lib/python3.7/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=0.9539842873176206, f1=0.9642857142857144,
best_f1=0.9697648376259798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07438772171735764
step: 10, loss: 0.00090721523156388959
step: 20, loss: 0.0012862246949225664
step: 30, loss: 0.0025295817758888006
step: 40, loss: 0.002329345792531967
step: 50, loss: 0.08847295492887497
step: 60, loss: 0.009739276021718979
step: 70, loss: 0.0019828276708722115
step: 80, loss: 0.005130224395543337
step: 90, loss: 0.000685464998241514
step: 100, loss: 0.0003142614441458136
step: 110, loss: 0.00824758131057024
step: 120, loss: 0.0062594059854745865
step: 130, loss: 0.0025472489651292562
step: 140, loss: 0.0007658626418560743
step: 150, loss: 0.00045716913882642984
step: 160, loss: 0.003549149027094245
step: 170, loss: 0.00044652799260802567
step: 180, loss: 0.0008115311502479017
step: 190, loss: 0.003547142492607236
step: 200, loss: 0.005184589885175228
step: 210, loss: 0.04486863687634468
step: 220, loss: 0.00026491336757317185
step: 230, loss: 0.0015318035148084164
/usr/local/lib/python3.7/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=0.9623893805309734, f1=0.9643652561247216,
best_f1=0.9697648376259798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1349784880876541
step: 10, loss: 0.0007777646533213556
step: 20, loss: 0.0002995934628415853
step: 30, loss: 0.020356804132461548
step: 40, loss: 0.07190675288438797
step: 50, loss: 0.0003381818241905421
step: 60, loss: 0.018533553928136826
step: 70, loss: 0.00018148882372770458
step: 80, loss: 0.00027834728825837374
step: 90, loss: 0.004722716752439737
step: 100, loss: 0.0023430536966770887
step: 110, loss: 0.0006838078261353076
step: 120, loss: 0.0016165263950824738
step: 130, loss: 0.0016464153304696083
step: 140, loss: 0.0006511714309453964
step: 150, loss: 0.0013781004818156362
step: 160, loss: 0.00019432419503573328
step: 170, loss: 0.004322740249335766
step: 180, loss: 0.00013622829283121973
step: 190, loss: 0.00031696693622507155
step: 200, loss: 0.00030609144596382976
step: 210, loss: 7.954976899782196e-05
step: 220, loss: 0.00027829851023852825
step: 230, loss: 0.06546780467033386
/usr/local/lib/python3.7/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=0.9547960308710033, f1=0.9599999999999999,
best_f1=0.9697648376259798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00566335953772068
step: 10, loss: 0.1104884222149849
step: 20, loss: 0.000962342950515449
step: 30, loss: 0.0027573436964303255
step: 40, loss: 0.0025220480747520924
step: 50, loss: 0.004083673004060984
step: 60, loss: 0.006604690104722977
step: 70, loss: 0.0006462483434006572
step: 80, loss: 0.0018244526581838727
step: 90, loss: 0.0002690728579182178
step: 100, loss: 0.0003705463605001569
step: 110, loss: 0.00023158255498856306
step: 120, loss: 0.00019000674365088344
step: 130, loss: 0.0007774244295433164
step: 140, loss: 0.0012026808690279722
```

```
step: 150, loss: 0.0003866699698846787
step: 160, loss: 0.00018981413450092077
step: 170, loss: 0.00028274027863517404
step: 180, loss: 0.003467332338914275
step: 190, loss: 0.00016431455151177943
step: 200, loss: 0.013945012353360653
step: 210, loss: 8.137624536175281e-05
step: 220, loss: 0.00016539578791707754
step: 230, loss: 0.13987991213798523
/usr/local/lib/python3.7/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=0.9587513935340021, f1=0.968609865470852,
best_f1=0.9697648376259798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007280217832885683
step: 10, loss: 0.00020058005629805848
step: 20, loss: 0.0008591603254899383
step: 30, loss: 0.0005114988307468593
step: 40, loss: 0.00017393774760421365
step: 50, loss: 0.012962386943399906
step: 60, loss: 0.005039168521761894
step: 70, loss: 0.0009400803246535361
step: 80, loss: 0.0010437587043270469
step: 90, loss: 0.001158394617959857
step: 100, loss: 0.0013699212577193975
step: 110, loss: 8.29819546197541e-05
step: 120, loss: 0.043488577008247375
step: 130, loss: 0.00017497364024166018
step: 140, loss: 0.01955287717282772
step: 150, loss: 7.048733095871285e-05
step: 160, loss: 0.0004580451059155166
step: 170, loss: 0.00011212395475013182
step: 180, loss: 0.013483071699738503
step: 190, loss: 0.0002656656433828175
step: 200, loss: 0.00022256755619309843
step: 210, loss: 0.00034242813126184046
step: 220, loss: 0.01170396339148283
step: 230, loss: 0.000265625974861905
/usr/local/lib/python3.7/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=0.9576837416481068, f1=0.9686800894854586,
best_f1=0.9697648376259798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011827965499833226
step: 10, loss: 9.521419269731268e-05
step: 20, loss: 9.900561417452991e-05
step: 30, loss: 0.0007159528904594481
step: 40, loss: 0.0006057330756448209
step: 50, loss: 0.0002710465923883021
step: 60, loss: 0.03628862649202347
step: 70, loss: 0.004992571193724871
step: 80, loss: 0.00032021154765971005
step: 90, loss: 0.0001542608079034835
step: 100, loss: 0.00012952083488926291
step: 110, loss: 0.013304587453603745
step: 120, loss: 0.01878293789923191
step: 130, loss: 0.0001463332591811195
step: 140, loss: 0.004958863370120525
step: 150, loss: 0.0006934304838068783
step: 160, loss: 0.00016558855713810772
step: 170, loss: 0.00015978662122506648
step: 180, loss: 0.08825114369392395
step: 190, loss: 0.00034712685737758875
step: 200, loss: 9.247838897863403e-05
step: 210, loss: 0.00011784653179347515
step: 220, loss: 0.003704913193359971
step: 230, loss: 0.0002935644588433206
/usr/local/lib/python3.7/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=0.9658213891951488, f1=0.9690949227373068,
best_f1=0.9697648376259798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016977431369014084
step: 10, loss: 0.0001156631187768653
step: 20, loss: 9.42553233471699e-05
step: 30, loss: 0.0003911591193173080 7
step: 40, loss: 0.10806410759687424
step: 50, loss: 0.0020881325472146273
step: 60, loss: 0.00041154242353513837
step: 70, loss: 0.014948730356991291
step: 80, loss: 0.000197296729311347
step: 90, loss: 0.01069028303027153
step: 100, loss: 0.0001278783311136067
step: 110, loss: 0.0002186162309953943
step: 120, loss: 0.00024671314167790115
step: 130, loss: 0.00012904478353448212
step: 140, loss: 0.0001250873610842973
```

```
step: 150, loss: 9.272791066905484e-05
step: 160, loss: 0.00011353470472386107
step: 170, loss: 0.005225012078881264
step: 180, loss: 0.00035249683423899114
step: 190, loss: 0.00019302564032841474
step: 200, loss: 9.777151717571542e-05
step: 210, loss: 7.357441063504666e-05
step: 220, loss: 0.00014748361718375236
step: 230, loss: 0.00021607505914289504
/usr/local/lib/python3.7/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=0.9633740288568259, f1=0.9678135405105438,
best_f1=0.9697648376259798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011200125300092623
step: 10, loss: 0.00011914942297153175
step: 20, loss: 6.916868005646393e-05
step: 30, loss: 0.00016126541595440358
step: 40, loss: 7.219032704597339e-05
step: 50, loss: 5.3569747251458466e-05
step: 60, loss: 5.7450444728601724e-05
step: 70, loss: 5.5986551160458475e-05
step: 80, loss: 5.966645403532311e-05
step: 90, loss: 0.00021241072681732476
step: 100, loss: 5.1099996198900044e-05
step: 110, loss: 0.0022698291577398777
step: 120, loss: 0.000112635811092332
step: 130, loss: 0.00040104106301441479
step: 140, loss: 5.199459337745793e-05
step: 150, loss: 9.017769480124116e-05
step: 160, loss: 0.00024207428214140236
step: 170, loss: 0.0001395273138768971
step: 180, loss: 7.66875091358088e-05
step: 190, loss: 0.00010349391959607601
step: 200, loss: 0.00048627337673678994
step: 210, loss: 9.07501089386642e-05
step: 220, loss: 0.007394622080028057
step: 230, loss: 0.002374125411733985
/usr/local/lib/python3.7/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=0.9594594594594594, f1=0.9685393258426966,
best_f1=0.9697648376259798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.151965549681336e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.00015600351616740227
step: 20, loss: 9.888599015539512e-05
step: 30, loss: 0.00010050613491330296
step: 40, loss: 0.00024526502238586545
step: 50, loss: 0.004207028541713953
step: 60, loss: 7.818442099960521e-05
step: 70, loss: 7.733359962003306e-05
step: 80, loss: 5.898550807614811e-05
step: 90, loss: 8.840488590067253e-05
step: 100, loss: 6.823783041909337e-05
step: 110, loss: 0.00019214156782254577
step: 120, loss: 9.16122953640297e-05
step: 130, loss: 0.06617631018161774
step: 140, loss: 0.036031804978847504
step: 150, loss: 0.00010127973655471578
step: 160, loss: 5.926500671193935e-05
step: 170, loss: 0.00010361718159401789
step: 180, loss: 0.00018903148884419352
step: 190, loss: 0.00012279313523322344
step: 200, loss: 0.00044648110633715987
step: 210, loss: 0.0008067029993981123
step: 220, loss: 0.0004128485161345452
step: 230, loss: 5.434119520941749e-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=0.966740576496674, f1=0.9700996677740864,
best_f1=0.9697648376259798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.86389204626903e-05
step: 10, loss: 4.643018473871052e-05
step: 20, loss: 0.0003272839530836791
step: 30, loss: 0.00017407507402822375
step: 40, loss: 7.519310747738928e-05
step: 50, loss: 0.0005517167737707496
step: 60, loss: 0.00029290124075487256
step: 70, loss: 8.262678602477536e-05
step: 80, loss: 0.00037615105975419283
step: 90, loss: 0.0001493066520197317
step: 100, loss: 0.0016326605109497905
step: 110, loss: 0.00026661725132726133
step: 120, loss: 8.613824320491403e-05
step: 130, loss: 0.00012443138984963298
```

```
step: 140, loss: 5.636350397253409e-05
step: 150, loss: 8.93571850610897e-05
step: 160, loss: 4.673148214351386e-05
step: 170, loss: 9.683071402832866e-05
step: 180, loss: 0.00019624341803137213
step: 190, loss: 5.317292016115971e-05
step: 200, loss: 6.556123844347894e-05
step: 210, loss: 0.00011510484182508662
step: 220, loss: 0.00018618974718265235
step: 230, loss: 4.347258436609991e-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=0.9655937846836848, f1=0.9678848283499446,
best_f1=0.9697648376259798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.156462768558413e-05
step: 10, loss: 0.00019308383343741298
step: 20, loss: 9.138697350863367e-05
step: 30, loss: 0.00011736483429558575
step: 40, loss: 0.00013291016512084752
step: 50, loss: 8.523935684934258e-05
step: 60, loss: 0.0002965585154015571
step: 70, loss: 0.00011685375648085028
step: 80, loss: 0.0008173899259418249
step: 90, loss: 3.989649485447444e-05
step: 100, loss: 0.00013325767940841615
step: 110, loss: 7.593930058646947e-05
step: 120, loss: 0.00019142779638059437
step: 130, loss: 9.918333671521395e-05
step: 140, loss: 0.009601221419870853
step: 150, loss: 0.00012338512169662863
step: 160, loss: 0.00033017934765666723
step: 170, loss: 4.97265973535832e-05
step: 180, loss: 0.00013071017747279257
step: 190, loss: 0.00020703602058347315
step: 200, loss: 7.536375778727233e-05
step: 210, loss: 0.0273995753377676
step: 220, loss: 0.0019419633317738771
step: 230, loss: 6.75954288453795e-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=0.9655172413793103, f1=0.9655937846836848,
best_f1=0.9697648376259798
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 [00:08, 301.77it/s]
load_f1 = 0.9655172413793103
real_f1 = 0.9666666666666666
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 [00:12, 348.28it/s]
```

1.3.2 DBLP-GoogleScholar - Running the matcher

```python
# 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 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: 0.8027960062026978
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.43144842982292175
step: 20, loss: 0.5000458359718323
step: 30, loss: 0.4503665566444397
step: 40, loss: 0.3904196619987488
step: 50, loss: 0.26284754276275635
step: 60, loss: 0.46936970949172974
step: 70, loss: 0.22096773982048035
step: 80, loss: 0.14481958746910095
step: 90, loss: 0.11053013801574707
step: 100, loss: 0.3165495991706848
step: 110, loss: 0.1327943354845047
step: 120, loss: 0.0753868892788887
step: 130, loss: 0.06279391050338745
step: 140, loss: 0.3629947900772095
step: 150, loss: 0.05931292846798897
step: 160, loss: 0.1743725687265396
step: 170, loss: 0.2603021264076233
step: 180, loss: 0.1534986048936844
step: 190, loss: 0.05716396123170853
step: 200, loss: 0.1139310896396637
step: 210, loss: 0.11849211156368256
step: 220, loss: 0.2762703001499176
step: 230, loss: 0.09481446444988251
step: 240, loss: 0.19095875322818756
step: 250, loss: 0.0417509600520134
step: 260, loss: 0.03651381656527519
step: 270, loss: 0.03517543524503708
step: 280, loss: 0.12436933070421219
step: 290, loss: 0.03846661001443863
step: 300, loss: 0.1714363396167755
step: 310, loss: 0.08191072195768356
step: 320, loss: 0.18863151967525482
step: 330, loss: 0.14083677530288696
```

```
step: 340, loss: 0.20301716029644012
step: 350, loss: 0.1619354635477066
step: 360, loss: 0.07333254814147949
step: 370, loss: 0.1525430530309677
step: 380, loss: 0.14184629917144775
step: 390, loss: 0.046260636299848557
step: 400, loss: 0.026958325877785683
step: 410, loss: 0.09045048803091049
step: 420, loss: 0.03885591775178909
step: 430, loss: 0.09244509041309357
step: 440, loss: 0.2223016768693924
step: 450, loss: 0.10597026348114014
step: 460, loss: 0.04652990773320198
step: 470, loss: 0.21474555134773254
step: 480, loss: 0.28861284255981445
step: 490, loss: 0.050921544432640076
step: 500, loss: 0.018425296992063522
step: 510, loss: 0.07406995445489883
step: 520, loss: 0.06361363083124161
step: 530, loss: 0.07001522928476334
/usr/local/lib/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=0.9153018249883015, f1=0.9130434782608695,
best_f1=0.9130434782608695
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15797103941440582
step: 10, loss: 0.17370346188545227
step: 20, loss: 0.13486117124557495
step: 30, loss: 0.09463094174861908
step: 40, loss: 0.024748194962739944
step: 50, loss: 0.05942172184586525
step: 60, loss: 0.21774077415466309
step: 70, loss: 0.0601564384996891
step: 80, loss: 0.13043507933616638
step: 90, loss: 0.06961669772863388
step: 100, loss: 0.3239063620567322
step: 110, loss: 0.04128827899694443
step: 120, loss: 0.2605676054954529
step: 130, loss: 0.12801344692707062
step: 140, loss: 0.16105607151985168
step: 150, loss: 0.04698958620429039
```

```
step: 160, loss: 0.02558242715895176
step: 170, loss: 0.1520395278930664
step: 180, loss: 0.007698520552366972
step: 190, loss: 0.07048577070236206
step: 200, loss: 0.045704107731580734
step: 210, loss: 0.07973587512969971
step: 220, loss: 0.11754093319177628
step: 230, loss: 0.08395878970623016
step: 240, loss: 0.11355547606945038
step: 250, loss: 0.045221682637929916
step: 260, loss: 0.022975606843829155
step: 270, loss: 0.06426086276769638
step: 280, loss: 0.2073928713798523
step: 290, loss: 0.03973238170146942
step: 300, loss: 0.10115768015384674
step: 310, loss: 0.032636579126119614
step: 320, loss: 0.18596471846103668
step: 330, loss: 0.08562503755092621
step: 340, loss: 0.0044715614058077335
step: 350, loss: 0.11804213374853134
step: 360, loss: 0.062107205390930176
step: 370, loss: 0.015543212182819843
step: 380, loss: 0.1194576695561409
step: 390, loss: 0.06200284883379936
step: 400, loss: 0.12947218120098114
step: 410, loss: 0.0013527680421248078
step: 420, loss: 0.05981529876589775
step: 430, loss: 0.014380138367414474
step: 440, loss: 0.03233388438820839
step: 450, loss: 0.022467415779829025
step: 460, loss: 0.2276000827550888
step: 470, loss: 0.031100330874323845
step: 480, loss: 0.105296291410923
step: 490, loss: 0.1908971667289734
step: 500, loss: 0.035726241767406464
step: 510, loss: 0.2133902758359909
step: 520, loss: 0.15943314135074615
step: 530, loss: 0.2817245125770569
/usr/local/lib/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=0.9291412482402628, f1=0.9174397031539888,
best_f1=0.9174397031539888
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008414430543780327
step: 10, loss: 0.16589076817035675
step: 20, loss: 0.3186394274234772
step: 30, loss: 0.18532724678516388
step: 40, loss: 0.026889193803071976
step: 50, loss: 0.027827240526676178
step: 60, loss: 0.00415390869602561
step: 70, loss: 0.022707561030983925
step: 80, loss: 0.14411576092243195
step: 90, loss: 0.14051592350006104
step: 100, loss: 0.06164904311299324
step: 110, loss: 0.026141062378883362
step: 120, loss: 0.036551207304000854
step: 130, loss: 0.009188024327158928
step: 140, loss: 0.07607022672891617
step: 150, loss: 0.021493714302778244
step: 160, loss: 0.003522695042192936
step: 170, loss: 0.15373288094997406
step: 180, loss: 0.016925489529967308
step: 190, loss: 0.01469370350241661
step: 200, loss: 0.027238689363002777
step: 210, loss: 0.08188199251890182
step: 220, loss: 0.031061183661222458
step: 230, loss: 0.006761852186173201
step: 240, loss: 0.016668720170855522
step: 250, loss: 0.013171187601983547
step: 260, loss: 0.018224462866783142
step: 270, loss: 0.007331632077693939
step: 280, loss: 0.17345434427261353
step: 290, loss: 0.06074219197034836
step: 300, loss: 0.14251287281513214
step: 310, loss: 0.20237720012664795
step: 320, loss: 0.13979415595531464
step: 330, loss: 0.015254185535013676
step: 340, loss: 0.012175485491752625
step: 350, loss: 0.050885964184999466
step: 360, loss: 0.016383366659283638
step: 370, loss: 0.004554336424916983
step: 380, loss: 0.0415339469909668
step: 390, loss: 0.009678098373115063
step: 400, loss: 0.022025030106306076
step: 410, loss: 0.027236230671405792
step: 420, loss: 0.08743225783109665
step: 430, loss: 0.09780857712030411
step: 440, loss: 0.017948629334568977
step: 450, loss: 0.018337713554501534
```

```
step: 460, loss: 0.06743913888931274
step: 470, loss: 0.00878389272838831
step: 480, loss: 0.014610477723181248
step: 490, loss: 0.09664341807365417
step: 500, loss: 0.10788407176733017
step: 510, loss: 0.0183674655854702
step: 520, loss: 0.008767754770815372
step: 530, loss: 0.018617454916238785
/usr/local/lib/python3.7/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=0.9200184928340269, f1=0.9254004576659038,
best_f1=0.9174397031539888
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005635957699269056
step: 10, loss: 0.020364627242088318
step: 20, loss: 0.017771124839782715
step: 30, loss: 0.05078008770942688
step: 40, loss: 0.09790783375501633
step: 50, loss: 0.07452841848134995
step: 60, loss: 0.035277318209409714
step: 70, loss: 0.00981142371892929
step: 80, loss: 0.102371945977211
step: 90, loss: 0.1623218059539795
step: 100, loss: 0.07098807394504547
step: 110, loss: 0.022370638325810432
step: 120, loss: 0.0007703949231654406
step: 130, loss: 0.07285014539957047
step: 140, loss: 0.049916550517082214
step: 150, loss: 0.007574915885925293
step: 160, loss: 0.004255667794495821
step: 170, loss: 0.027624215930700302
step: 180, loss: 0.0725846141576767
step: 190, loss: 0.008775921538472176
step: 200, loss: 0.001066525699570775
step: 210, loss: 0.078124038875103
step: 220, loss: 0.011879196390509605
step: 230, loss: 0.22481174767017365
step: 240, loss: 0.002856332575902343
step: 250, loss: 0.013366390950977802
step: 260, loss: 0.13309578597545624
step: 270, loss: 0.09873639792203903
step: 280, loss: 0.008331610821187496
step: 290, loss: 0.09079382568597794
step: 300, loss: 0.012448419816792011
step: 310, loss: 0.006727650761604309
```

```
step: 320, loss: 0.13349223136901855
step: 330, loss: 0.044806886464357376
step: 340, loss: 0.04025767743587494
step: 350, loss: 0.0033135879784822464
step: 360, loss: 0.015266237780451775
step: 370, loss: 0.13441264629364014
step: 380, loss: 0.007042403798550367
step: 390, loss: 0.009472953155636787
step: 400, loss: 0.045036427676677704
step: 410, loss: 0.004635002929717302
step: 420, loss: 0.013195300474762917
step: 430, loss: 0.023947346955537796
step: 440, loss: 0.05050178989768028
step: 450, loss: 0.013091789558529854
step: 460, loss: 0.048396192491054535
step: 470, loss: 0.003001183969900012
step: 480, loss: 0.07820083200931549
step: 490, loss: 0.05892713740468025
step: 500, loss: 0.022490808740258217
step: 510, loss: 0.058833151564002037
step: 520, loss: 0.08944739401340485
step: 530, loss: 0.0012341142864897847
/usr/local/lib/python3.7/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=0.924791086350975, f1=0.9220839096357768,
best_f1=0.9174397031539888
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0047870175912976265
step: 10, loss: 0.06900651752948761
step: 20, loss: 0.0011125242793932557
step: 30, loss: 0.00193695945199579
step: 40, loss: 0.06619421392679214
step: 50, loss: 0.019522899761795998
step: 60, loss: 0.0017994135851040483
step: 70, loss: 0.0031533888541162014
step: 80, loss: 0.004213653039187193
step: 90, loss: 0.03330003842711449
step: 100, loss: 0.0007593509508296847
step: 110, loss: 0.0008356309845112264
step: 120, loss: 0.0696396753191948
step: 130, loss: 0.0024777408689260483
step: 140, loss: 0.004135516472160816
step: 150, loss: 0.0017150405328720808
step: 160, loss: 0.002106839558109641
step: 170, loss: 0.0009073616820387542
```

```
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 180, loss: 0.0006036291015334427
step: 190, loss: 0.0008499221294187009
step: 200, loss: 0.05388268083333969
step: 210, loss: 0.003795723430812359
step: 220, loss: 0.015773996710777283
step: 230, loss: 0.0018205327214673162
step: 240, loss: 0.08674998581409454
step: 250, loss: 0.003901570802554488
step: 260, loss: 0.012465653009712696
step: 270, loss: 0.050139956176280975
step: 280, loss: 0.0490933433175087
step: 290, loss: 0.019535619765520096
step: 300, loss: 0.03802407160401344
step: 310, loss: 0.0024138360749930143
step: 320, loss: 0.013678213581442833
step: 330, loss: 0.00916708167642355
step: 340, loss: 0.009711433202028275
step: 350, loss: 0.015805918723344803
step: 360, loss: 0.002394256414845586
step: 370, loss: 0.05570443719625473
step: 380, loss: 0.04480989649891853
step: 390, loss: 0.005182545632123947
step: 400, loss: 0.03114665299654007
step: 410, loss: 0.000826189061626792
step: 420, loss: 0.0018601008923724294
step: 430, loss: 0.044650670140981674
step: 440, loss: 0.008782299235463142
step: 450, loss: 0.03483636677265167
step: 460, loss: 0.006251832935959101
step: 470, loss: 0.0022642137482762337
step: 480, loss: 0.004355115815997124
step: 490, loss: 0.014006288722157478
step: 500, loss: 0.005468698684126139
step: 510, loss: 0.022006070241332054
step: 520, loss: 0.0011238910956308246
step: 530, loss: 0.004550617653876543
/usr/local/lib/python3.7/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=0.9249999999999999, f1=0.9239981575310916,
best_f1=0.9174397031539888
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006856819614768028
step: 10, loss: 0.0036206552758812904
step: 20, loss: 0.0009287084685638547
```

```
step: 30, loss: 0.004635270684957504
step: 40, loss: 0.13709959387779236
step: 50, loss: 0.015489870682358742
step: 60, loss: 0.006504382938146591
step: 70, loss: 0.053222235292196274
step: 80, loss: 0.008101758547127247
step: 90, loss: 0.013963333331048489
step: 100, loss: 0.060231853276491165
step: 110, loss: 0.014935294166207314
step: 120, loss: 0.0538700595498085
step: 130, loss: 0.046141937375068665
step: 140, loss: 0.0044332570396363735
step: 150, loss: 0.031624093651771545
step: 160, loss: 0.0007851515547372401
step: 170, loss: 0.00041203672299161553
step: 180, loss: 0.07665251195430756
step: 190, loss: 0.06226031854748726
step: 200, loss: 0.0006433050730265677
step: 210, loss: 0.01600552722811699
step: 220, loss: 0.0034733284264802933
step: 230, loss: 0.0002594898105598986
step: 240, loss: 0.01902684010565281
step: 250, loss: 0.0027819545939564705
step: 260, loss: 0.011643754318356514
step: 270, loss: 0.0016219503013417125
step: 280, loss: 0.03707215189933777
step: 290, loss: 0.11035362631082535
step: 300, loss: 0.0046175336465239525
step: 310, loss: 0.007939798757433891
step: 320, loss: 0.015211524441838264
step: 330, loss: 0.09041857719421387
step: 340, loss: 0.007129184901714325
step: 350, loss: 0.007011254318058491
step: 360, loss: 0.010978898033499718
step: 370, loss: 0.15437236428260803
step: 380, loss: 0.05765507370233536
step: 390, loss: 0.013501053676009178
step: 400, loss: 0.025985611602663994
step: 410, loss: 0.0007627259474247694
step: 420, loss: 0.0242879968136549
step: 430, loss: 0.0030867864843457937
step: 440, loss: 0.01790589839220047
step: 450, loss: 0.0037544025108218193
step: 460, loss: 0.005047918763011694
step: 470, loss: 0.021357180550694466
step: 480, loss: 0.011736548505723476
step: 490, loss: 0.00299626961350441
step: 500, loss: 0.0013748290948569775
```

```
step: 510, loss: 0.000464823649963364
step: 520, loss: 0.012268871068954468
step: 530, loss: 0.11940965801477432
/usr/local/lib/python3.7/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=0.9135687732342007, f1=0.9204440333024977,
best_f1=0.9174397031539888
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008582398295402527
step: 10, loss: 0.017139248549938202
step: 20, loss: 0.0044952272437512875
step: 30, loss: 0.13681741058826447
step: 40, loss: 0.00962843932211399
step: 50, loss: 0.06924690306186676
step: 60, loss: 0.12550900876522064
step: 70, loss: 0.007679736707359552
step: 80, loss: 0.02955368161201477
step: 90, loss: 0.023089531809091568
step: 100, loss: 0.0012654946185648441
step: 110, loss: 0.011526601389050484
step: 120, loss: 0.001364945899695158
step: 130, loss: 0.00246029207482934
step: 140, loss: 0.0076422239653766155
step: 150, loss: 0.006773826200515032
step: 160, loss: 0.0011054497445002198
step: 170, loss: 0.004634690936654806
step: 180, loss: 0.0005071120103821158
step: 190, loss: 0.0001340161106782034
step: 200, loss: 0.02883717603981495
step: 210, loss: 0.005017990246415138
step: 220, loss: 0.0005413953331299126
step: 230, loss: 0.0019531301222741604
step: 240, loss: 0.0013706318568438292
step: 250, loss: 0.002997372532263398
step: 260, loss: 0.0006191371940076351
step: 270, loss: 0.00036684522638097405
step: 280, loss: 0.14098672568798065
step: 290, loss: 0.029410380870103836
step: 300, loss: 0.000491882034111768
step: 310, loss: 0.005036667454987764
step: 320, loss: 0.041324432939291
step: 330, loss: 0.0005525163724087179
step: 340, loss: 0.004465806297957897
step: 350, loss: 0.05201514810323715
step: 360, loss: 0.004213119391351938
```

```
step: 370, loss: 0.0017641213489696383
step: 380, loss: 0.0019939348567277193
step: 390, loss: 0.0021483844611793756
step: 400, loss: 0.006276419386267662
step: 410, loss: 0.005256091710180044
step: 420, loss: 0.0008132929215207696
step: 430, loss: 0.00010627588926581666
step: 440, loss: 0.00218567019328475
step: 450, loss: 0.0008158425916917622
step: 460, loss: 0.008730234578251839
step: 470, loss: 0.0038138485979288816
step: 480, loss: 0.0006793395150452852
step: 490, loss: 0.0004360961611382663
step: 500, loss: 0.00020345242228358984
step: 510, loss: 0.002895480254665017
step: 520, loss: 0.00107811507768929
step: 530, loss: 0.0002927168970927596
/usr/local/lib/python3.7/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=0.9225058004640372, f1=0.9279112754158966,
best_f1=0.9174397031539888
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012469443492591381
step: 10, loss: 0.008681447245180607
step: 20, loss: 0.024250606074929237
step: 30, loss: 0.023157529532909393
step: 40, loss: 0.0013045375235378742
step: 50, loss: 0.00022844967315904796
step: 60, loss: 0.08585460484027863
step: 70, loss: 0.014433744363486767
step: 80, loss: 0.000127939842059277
step: 90, loss: 0.04558301344513893
step: 100, loss: 0.004893863573670387
step: 110, loss: 0.00437801331281662
step: 120, loss: 0.0034512728452682495
step: 130, loss: 0.0005006769788451493
step: 140, loss: 0.00011377965711290017
step: 150, loss: 0.0029276316054165363
step: 160, loss: 0.0011797903571277857
step: 170, loss: 0.014300153590738773
step: 180, loss: 0.0005971876089461148
step: 190, loss: 0.0010207396699115634
step: 200, loss: 0.0037495240103453398
step: 210, loss: 0.001809217850677669
step: 220, loss: 0.002594747580587864
```

```
step: 230, loss: 0.012065110728144646
step: 240, loss: 0.2302277386188507
step: 250, loss: 0.05821676924824715
step: 260, loss: 0.07316748797893524
step: 270, loss: 0.0005555120878852904
step: 280, loss: 0.011877397075295448
step: 290, loss: 0.0021092905662953854
step: 300, loss: 0.00041539527592249215
step: 310, loss: 0.04580157622694969
step: 320, loss: 0.011369960382580757
step: 330, loss: 0.04041275754570961
step: 340, loss: 0.0021380498073995113
step: 350, loss: 0.0037969332188367844
step: 360, loss: 0.001798430341295898
step: 370, loss: 0.0012419353006407619
step: 380, loss: 0.054163578897714615
step: 390, loss: 0.0020357081666588783
step: 400, loss: 0.020206205546855927
step: 410, loss: 0.015935344621539116
step: 420, loss: 0.0005104460869915783
step: 430, loss: 0.00047821024782024324
step: 440, loss: 0.0007911105640232563
step: 450, loss: 0.0010799579322338104
step: 460, loss: 0.0010569972218945622
step: 470, loss: 0.006855905055999756
step: 480, loss: 0.00024289537395816296
step: 490, loss: 0.01838632859289646
step: 500, loss: 0.0038773922715336084
step: 510, loss: 0.0003462318563833833
step: 520, loss: 0.0008580168941989541
step: 530, loss: 0.00022633952903561294
/usr/local/lib/python3.7/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=0.9213377296278851, f1=0.9211385907606159,
best_f1=0.9174397031539888
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010498732328414917
step: 10, loss: 0.0017478149384260178
step: 20, loss: 0.0012187763350084424
step: 30, loss: 0.0024257346522063017
step: 40, loss: 0.01880977302789688
step: 50, loss: 0.002751419087871909
step: 60, loss: 5.900311953155324e-05
step: 70, loss: 0.11716799437999725
step: 80, loss: 0.0050529250875115395
```

```
step: 90, loss: 0.0036980637814849615
step: 100, loss: 0.00048202023026533425
step: 110, loss: 0.0060429079458117485
step: 120, loss: 0.0008298041648231447
step: 130, loss: 0.0007641517440788448
step: 140, loss: 0.10041213035583496
step: 150, loss: 0.0021486482582986355
step: 160, loss: 0.0022000723984092474
step: 170, loss: 0.00430749497562647
step: 180, loss: 0.000187701967661269
step: 190, loss: 0.0009319486562162638
step: 200, loss: 0.0032702283933758736
step: 210, loss: 0.0008292914135381579
step: 220, loss: 0.005651275161653757
step: 230, loss: 0.004837235901504755
step: 240, loss: 0.001161943539045751
step: 250, loss: 0.012075540609657764
step: 260, loss: 0.004981083329766989
step: 270, loss: 0.010615292936563492
step: 280, loss: 0.00025205471320077777
step: 290, loss: 0.0011131312930956483
step: 300, loss: 0.0009452279773540795
step: 310, loss: 0.00017164101882372051
step: 320, loss: 0.0003416461986489594
step: 330, loss: 0.0035088430158793926
step: 340, loss: 0.0002621675084810704
step: 350, loss: 0.00016604273696430027
step: 360, loss: 0.04187082126736641
step: 370, loss: 0.0016546188853681087
step: 380, loss: 8.201643504435197e-05
step: 390, loss: 0.0007925948593765497
step: 400, loss: 0.21565327048301697
step: 410, loss: 0.006197983864694834
step: 420, loss: 0.004736088681966066
step: 430, loss: 0.00012265381519682705
step: 440, loss: 0.003081786911934614
step: 450, loss: 0.0025627161376178265
step: 460, loss: 0.0013667354360222816
step: 470, loss: 0.00032214567181654274
step: 480, loss: 0.003375914879143238
step: 490, loss: 0.0019652405753731728
step: 500, loss: 0.004320580046623945
step: 510, loss: 0.0007278809789568186
step: 520, loss: 0.0013454772997647524
step: 530, loss: 0.0005313360597938299
/usr/local/lib/python3.7/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=0.9241639189825719, f1=0.9180327868852459,
best_f1=0.9174397031539888
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013076047180220485
step: 10, loss: 0.0006766266888007522
step: 20, loss: 0.0003236025804653764
step: 30, loss: 0.0014598607085645199
step: 40, loss: 0.0006358372629620135
step: 50, loss: 0.0006141425110399723
step: 60, loss: 0.0009553380077704787
step: 70, loss: 0.0025227696169167757
step: 80, loss: 0.002225494012236595
step: 90, loss: 0.04658294469118118
step: 100, loss: 0.015747200697660446
step: 110, loss: 0.0013029887340962887
step: 120, loss: 0.0009096486028283834
step: 130, loss: 0.00048325557145290077
step: 140, loss: 0.003988336771726608
step: 150, loss: 0.004335507750511169
step: 160, loss: 0.0001872172870207578
step: 170, loss: 0.0020276308059692383
step: 180, loss: 0.010565032251179218
step: 190, loss: 0.0054022520780563354
step: 200, loss: 0.0005865185521543026
step: 210, loss: 0.0005971150239929557
step: 220, loss: 0.00011830015137093142
step: 230, loss: 0.00023070347378961742
step: 240, loss: 0.0014501389814540744
step: 250, loss: 0.0048760101199150085
step: 260, loss: 0.012244713492691517
step: 270, loss: 6.620182830374688e-05
step: 280, loss: 3.3769014407880604e-05
step: 290, loss: 0.0004691685608122498
step: 300, loss: 0.002492383122444153
step: 310, loss: 0.00019758119015023112
step: 320, loss: 0.0007975054322741926
step: 330, loss: 0.0009423668961971998
step: 340, loss: 0.00014789328270126134
step: 350, loss: 0.01275810319930315
step: 360, loss: 6.5671767515596e-05
step: 370, loss: 0.009267783723771572
step: 380, loss: 0.0002665890206117183
step: 390, loss: 0.00011697147419909015
step: 400, loss: 0.01387642789632082
step: 410, loss: 0.0001892834552563727
step: 420, loss: 0.0053948708809912205
```

```
step: 430, loss: 0.000190304868738167
step: 440, loss: 0.00011311105481581762
step: 450, loss: 0.001829107291996479
step: 460, loss: 0.002679014578461647
step: 470, loss: 0.00023826980032026768
step: 480, loss: 0.0006749754538759589
step: 490, loss: 0.09245181828737259
step: 500, loss: 0.08628799766302109
step: 510, loss: 0.00462966738268733
step: 520, loss: 0.007214718032628298
step: 530, loss: 6.35274191154167e-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=0.9185457892314773, f1=0.9200726612170754,
best_f1=0.9174397031539888
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006401054561138153
step: 10, loss: 0.0020277639850974083
step: 20, loss: 0.0007495504687540233
step: 30, loss: 0.00018095702398568392
step: 40, loss: 0.006507003679871559
step: 50, loss: 0.023186443373560905
step: 60, loss: 6.697475328110158e-05
step: 70, loss: 0.00022280245320871472
step: 80, loss: 0.00043610678403638303
step: 90, loss: 0.011441904120147228
step: 100, loss: 0.0014912558253854513
step: 110, loss: 6.290626333793625e-05
step: 120, loss: 0.0018187680980190635
step: 130, loss: 0.00041467594564892235
step: 140, loss: 0.0006389656336978078
step: 150, loss: 0.00064039824064821
step: 160, loss: 0.006588354241102934
step: 170, loss: 0.0006557541200891137
step: 180, loss: 6.697628123220056e-05
step: 190, loss: 0.000963917060289532
step: 200, loss: 0.0015851486241444945
step: 210, loss: 0.003647554200142622
step: 220, loss: 0.002885547000914812
step: 230, loss: 0.001377301407046616
step: 240, loss: 4.117770731681958e-05
step: 250, loss: 0.00011458367225714028
step: 260, loss: 0.0004949768190272152
step: 270, loss: 0.004183621611446142
step: 280, loss: 0.009775024838745594
```

```
step: 290, loss: 0.0032685548067092896
step: 300, loss: 0.011804039590060711
step: 310, loss: 0.026056483387947083
step: 320, loss: 0.0007813236443325877
step: 330, loss: 0.0006820610142312944
step: 340, loss: 4.801990507985465e-05
step: 350, loss: 0.00020157487597316504
step: 360, loss: 0.011285156942903996
step: 370, loss: 0.0006580625777132809
step: 380, loss: 0.0002967508335132152
step: 390, loss: 0.019831260666251183
step: 400, loss: 0.00010798485891427845
step: 410, loss: 0.00021802762057632208
step: 420, loss: 0.00038369424873963
step: 430, loss: 0.005784882698208094
step: 440, loss: 0.00012218535994179547
step: 450, loss: 0.0012568372767418623
step: 460, loss: 0.008404226042330265
step: 470, loss: 0.0009917003335431218
step: 480, loss: 0.10308892279863358
step: 490, loss: 0.03222573176026344
step: 500, loss: 0.002483749995008111
step: 510, loss: 0.0007917176699265838
step: 520, loss: 0.00018461071886122227
step: 530, loss: 0.00068667036248371
/usr/local/lib/python3.7/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=0.919815668202765, f1=0.9224452554744527,
best_f1=0.9174397031539888
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013660520780831575
step: 10, loss: 0.001114961109124124
step: 20, loss: 0.00010960753570543602
step: 30, loss: 0.005618019960820675
step: 40, loss: 0.0007008311804383993
step: 50, loss: 0.0011548524489626288
step: 60, loss: 0.022747410461306572
step: 70, loss: 0.0031474691350013018
step: 80, loss: 0.0020645069889724255
step: 90, loss: 0.0014073447091504931
step: 100, loss: 0.003303063102066517
step: 110, loss: 0.0004986742278560996
step: 120, loss: 0.000254072219831869
step: 130, loss: 0.0008160998113453388
step: 140, loss: 0.00021053206000942737
```

```
step: 150, loss: 0.0013936967588961124
step: 160, loss: 0.0017893476178869605
step: 170, loss: 0.007393438834697008
step: 180, loss: 0.004363496787846088
step: 190, loss: 0.000617321755271405
step: 200, loss: 0.0023619525600224733
step: 210, loss: 0.002193315187469125
step: 220, loss: 0.00012523549958132207
step: 230, loss: 0.0007080709328874946
step: 240, loss: 0.00013647768355440348
step: 250, loss: 0.00021947758796159178
step: 260, loss: 5.535614400287159e-05
step: 270, loss: 5.908307866775431e-05
step: 280, loss: 0.1049615815281868
step: 290, loss: 0.019802410155534744
step: 300, loss: 9.234061872120947e-05
step: 310, loss: 0.0005173183744773269
step: 320, loss: 0.00046752035268582404
step: 330, loss: 0.00023643542954232544
step: 340, loss: 0.0006869734497740865
step: 350, loss: 0.021024830639362335
step: 360, loss: 0.011958900839090347
step: 370, loss: 0.001435874029994011
step: 380, loss: 0.00018675978935789317
step: 390, loss: 6.127310916781425e-05
step: 400, loss: 3.9314531022682786e-05
step: 410, loss: 0.006577145773917437
step: 420, loss: 0.0012258344795554876
step: 430, loss: 0.0015906818443909287
step: 440, loss: 0.0003325273282825947
step: 450, loss: 0.016565212979912758
step: 460, loss: 0.0007228878093883395
step: 470, loss: 3.684787225211039e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 480, loss: 0.00023002614034339786
step: 490, loss: 0.0006991549162194133
step: 500, loss: 0.001662970520555973
step: 510, loss: 0.002229145262390375
step: 520, loss: 0.0003117398591712117
step: 530, loss: 0.0004952301387675107
/usr/local/lib/python3.7/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=0.9212410501193318, f1=0.9178147268408551,
best_f1=0.9174397031539888
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.0020413361489772797
step: 10, loss: 2.638208752614446e-05
step: 20, loss: 0.013969912193715572
step: 30, loss: 0.12141244858503342
step: 40, loss: 5.5448603234253824e-05
step: 50, loss: 0.0002226239157607779
step: 60, loss: 0.004844223149120808
step: 70, loss: 0.023718366399407387
step: 80, loss: 0.06045447289943695
step: 90, loss: 0.0016816298011690378
step: 100, loss: 0.00010452412243466824
step: 110, loss: 0.0021887796465307474
step: 120, loss: 0.00019692211935762316
step: 130, loss: 0.0017724265344440937
step: 140, loss: 0.00034511348349042237
step: 150, loss: 0.0015866872854530811
step: 160, loss: 7.449257100233808e-05
step: 170, loss: 0.0002747538674157113
step: 180, loss: 6.380239210557193e-05
step: 190, loss: 0.00015205214731395245
step: 200, loss: 0.0015312996692955494
step: 210, loss: 0.0014772447757422924
step: 220, loss: 0.00042456091614440084
step: 230, loss: 0.0001427404349669814
step: 240, loss: 7.698864646954462e-05
step: 250, loss: 0.04222607985138893
step: 260, loss: 0.009019184857606888
step: 270, loss: 0.0017550940392538905
step: 280, loss: 0.0028025535866618156
step: 290, loss: 0.0002260773762827739
step: 300, loss: 0.0005505536682903767
step: 310, loss: 5.772071745013818e-05
step: 320, loss: 6.1875740357209e-05
step: 330, loss: 0.0008587944903410971
step: 340, loss: 0.0003557071031536907
step: 350, loss: 0.0075590768828988075
step: 360, loss: 0.00013173703337088227
step: 370, loss: 0.0035823769867420197
step: 380, loss: 0.0014253025874495506
step: 390, loss: 0.00020029865845572203
step: 400, loss: 4.9143898650072515e-05
step: 410, loss: 5.938177491771057e-05
step: 420, loss: 0.0007429496035911143
step: 430, loss: 0.18803724646568298
step: 440, loss: 0.00270617357455194
step: 450, loss: 2.929116271843668e-05
step: 460, loss: 2.7688394766300917e-05
step: 470, loss: 0.02430332638323307
```

```
step: 480, loss: 0.0031525357626378536
step: 490, loss: 0.0020583143923431635
step: 500, loss: 0.00017786354874260724
step: 510, loss: 0.00022713841462973505
step: 520, loss: 0.0001813113340176642
step: 530, loss: 0.0002878776576835662
/usr/local/lib/python3.7/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=0.9238770685579196, f1=0.9231490159325211,
best_f1=0.9174397031539888
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007221633568406105
step: 10, loss: 0.0007741812150925398
step: 20, loss: 0.00024392874911427498
step: 30, loss: 0.0014377294573932886
step: 40, loss: 0.00014979616389609873
step: 50, loss: 0.0007309388020075858
step: 60, loss: 2.6694635380408727e-05
step: 70, loss: 8.82835010997951e-05
step: 80, loss: 5.4835170885780826e-05
step: 90, loss: 7.836199074517936e-05
step: 100, loss: 0.0003817401884589344
step: 110, loss: 0.00022930564591661096
step: 120, loss: 0.0026102170813828707
step: 130, loss: 0.0008157154661603272
step: 140, loss: 0.010133333504199982
step: 150, loss: 0.0008909471798688173
step: 160, loss: 0.0011044539278373122
step: 170, loss: 0.007221931591629982
step: 180, loss: 0.0693921446800232
step: 190, loss: 0.000283546483842656
step: 200, loss: 0.00028161591035313904
step: 210, loss: 0.014216174371540546
step: 220, loss: 0.0010480574565008283
step: 230, loss: 0.0013233608333393931
step: 240, loss: 0.003998038824647665
step: 250, loss: 0.003578159026801586
step: 260, loss: 0.00019189200247637928
step: 270, loss: 0.0046103717759251595
step: 280, loss: 0.0005233656847849488
step: 290, loss: 0.0014788864646106958
step: 300, loss: 0.0003818557306658476
step: 310, loss: 3.511649265419692e-05
step: 320, loss: 0.002295043086633086
step: 330, loss: 0.0005389477591961622
```

```
step: 340, loss: 0.0001262742734979838
step: 350, loss: 0.0013048930559307337
step: 360, loss: 0.0018762778490781784
step: 370, loss: 0.000941694073844701
step: 380, loss: 0.00015494275430683047
step: 390, loss: 0.0008810429135337472
step: 400, loss: 2.1110863599460572e-05
step: 410, loss: 0.0014858925715088844
step: 420, loss: 8.141342550516129e-05
step: 430, loss: 9.763612615643069e-05
step: 440, loss: 0.0015835327794775367
step: 450, loss: 0.0003380590642336756
step: 460, loss: 0.001728269038721919
step: 470, loss: 0.003303876146674156
step: 480, loss: 0.000849796982947737
step: 490, loss: 0.0004947853158228099
step: 500, loss: 0.0017005131812766194
step: 510, loss: 0.0025395252741873264
step: 520, loss: 0.00140548893250525
step: 530, loss: 2.6564237487036735e-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=0.9227144203581528, f1=0.9268065268065269,
best_f1=0.9174397031539888
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00038343097548931837
step: 10, loss: 0.00017862529784906656
step: 20, loss: 0.00010883057257160544
step: 30, loss: 0.0011230310192331672
step: 40, loss: 3.985797593486495e-05
step: 50, loss: 0.0014999380800873041
step: 60, loss: 0.0031884885393083096
step: 70, loss: 0.00023170067288447171
step: 80, loss: 4.7720335714984685e-05
step: 90, loss: 3.080923852394335e-05
step: 100, loss: 0.000491477083414793
step: 110, loss: 0.08904898166656494
step: 120, loss: 0.0011212273966521025
step: 130, loss: 4.829762110603042e-05
step: 140, loss: 2.3993707145564258e-05
step: 150, loss: 0.005471852142363787
step: 160, loss: 0.005318864248692989
step: 170, loss: 0.00013352402311284095
step: 180, loss: 0.00012548452650662512
step: 190, loss: 9.561661863699555e-05
```

```
step: 200, loss: 4.556850763037801e-05
step: 210, loss: 0.03654182702302933
step: 220, loss: 0.0001718963321764022
step: 230, loss: 0.0004906131071038544
step: 240, loss: 0.0004131157766096294
step: 250, loss: 8.052844350459054e-05
step: 260, loss: 0.0013862806372344494
step: 270, loss: 0.009400581009685993
step: 280, loss: 0.0026141314301639795
step: 290, loss: 0.001222087419591844
step: 300, loss: 0.23085390031337738
step: 310, loss: 0.0003795342636294663
step: 320, loss: 0.0004050004354212433
step: 330, loss: 0.0004339668666943908
step: 340, loss: 0.0010404433123767376
step: 350, loss: 0.00015017438272479922
step: 360, loss: 0.0005232459516264498
step: 370, loss: 0.00369943561963737
step: 380, loss: 0.0011816357728093863
step: 390, loss: 0.0002173283719457686
step: 400, loss: 0.0007579228258691728
step: 410, loss: 0.0019311538198962808
step: 420, loss: 9.644734382163733e-05
step: 430, loss: 3.9092421502573416e-05
step: 440, loss: 0.0004623572458513081
step: 450, loss: 0.0017929567256942391
step: 460, loss: 4.676339085563086e-05
step: 470, loss: 0.0002533882507123053
step: 480, loss: 0.00033641481422819197
step: 490, loss: 0.0004644864529836923
step: 500, loss: 0.01723080314695835
step: 510, loss: 0.0035672569647431374
step: 520, loss: 0.00013871886767446995
step: 530, loss: 6.571879202965647e-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=0.9227144203581528, f1=0.9247311827956989,
best_f1=0.9174397031539888
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.")
5742it [00:16, 340.06it/s]
load_f1 = 0.92524682651622
real_f1 = 0.9224952741020794
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 [00:12, 365.77it/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 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: 0.8672758936882019
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=0.2692307692307693, f1=0.2666666666666667,
best_f1=0.2666666666666667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.35362130403518677
/usr/local/lib/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=0.34615384615384615, f1=0.37681159420289856,
best_f1=0.37681159420289856
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3115081787109375
/usr/local/lib/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=0.43902439024390244, f1=0.38596491228070184,
best_f1=0.38596491228070184
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31752121448516846

```
/usr/local/lib/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=0.5333333333333333, f1=0.41025641025641024,
best_f1=0.41025641025641024
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19741153717041016
/usr/local/lib/python3.7/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=0.5333333333333333, f1=0.4, best_f1=0.41025641025641024
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25818148255348206
/usr/local/lib/python3.7/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=0.5142857142857143, f1=0.4680851063829786,
best_f1=0.41025641025641024
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16932553052902222
/usr/local/lib/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=0.5625000000000001, f1=0.43478260869565216,
best_f1=0.43478260869565216
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30693328380584717
/usr/local/lib/python3.7/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=0.5517241379310344, f1=0.45, best_f1=0.43478260869565216
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17311719059944153
/usr/local/lib/python3.7/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=0.5625000000000001, f1=0.4761904761904762,
best_f1=0.43478260869565216
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2127400040626526
/usr/local/lib/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=0.5806451612903226, f1=0.5128205128205129,
best_f1=0.5128205128205129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21485891938209534
/usr/local/lib/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=0.5925925925925927, f1=0.48484848484848486,
best_f1=0.48484848484848486
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12506243586540222
/usr/local/lib/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=0.6153846153846153, f1=0.5625000000000001,
best_f1=0.5625000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11566881835460663
/usr/local/lib/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=0.64, f1=0.5625000000000001, best_f1=0.5625000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1679956316947937
/usr/local/lib/python3.7/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=0.64, f1=0.5625000000000001, best_f1=0.5625000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.188396155834198
/usr/local/lib/python3.7/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=0.64, f1=0.5625000000000001, best_f1=0.5625000000000001
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, 121865.15it/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 = 0.64
real_f1 = 0.64
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, 256.87it/s]
```

2.1.2 DBLP-ACM - Running the matcher

```python
[ ]: # 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/
```

140

```
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: 0.8159424662590027
step: 10, loss: 0.46650299429893494
step: 20, loss: 0.5871344804763794
step: 30, loss: 0.45518165826797485
step: 40, loss: 0.26987922191619873
step: 50, loss: 0.136795774102211
step: 60, loss: 0.1262657344341278
step: 70, loss: 0.10253632813692093
step: 80, loss: 0.22642461955547333
step: 90, loss: 0.01664246991276741
step: 100, loss: 0.15998516976833344
step: 110, loss: 0.03364115580916405
```

```
step: 120, loss: 0.10695276409387589
step: 130, loss: 0.0038966061547398567
step: 140, loss: 0.06976217776536942
step: 150, loss: 0.01790049485862255
step: 160, loss: 0.20987704396247864
step: 170, loss: 0.005916181020438671
step: 180, loss: 0.013844043016433716
step: 190, loss: 0.02718304842710495
step: 200, loss: 0.0042565492913126945
step: 210, loss: 0.004780846647918224
step: 220, loss: 0.009659981355071068
step: 230, loss: 0.015394027344882488
/usr/local/lib/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=0.9642058165548099, f1=0.9626274065685164,
best_f1=0.9626274065685164
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03780259191989899
step: 10, loss: 0.0022198499646037817
step: 20, loss: 0.003331694984808564
step: 30, loss: 0.010601687245070934
step: 40, loss: 0.008928358554840088
step: 50, loss: 0.03916093707084656
step: 60, loss: 0.0024236866738647223
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 70, loss: 0.004998985677957535
step: 80, loss: 0.10275259613990784
step: 90, loss: 0.00437694787979126
step: 100, loss: 0.021401308476924896
step: 110, loss: 0.002086317865177989
step: 120, loss: 0.00466473214328289
step: 130, loss: 0.006705678533762693
step: 140, loss: 0.038657668977975845
step: 150, loss: 0.015006710775196552
step: 160, loss: 0.007030919659882784
step: 170, loss: 0.11848524212837219
step: 180, loss: 0.0031230978202074766
step: 190, loss: 0.09542307257652283
step: 200, loss: 0.004514478612691164
step: 210, loss: 0.11313888430595398
step: 220, loss: 0.0020844913087785244
```

```
step: 230, loss: 0.2468469738960266
/usr/local/lib/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=0.9807474518686297, 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: 0.056313276290893555
step: 10, loss: 0.016076017171144485
step: 20, loss: 0.011452237144112587
step: 30, loss: 0.16050684452056885
step: 40, loss: 0.04618493467569351
step: 50, loss: 0.005188464652746916
step: 60, loss: 0.010962331667542458
step: 70, loss: 0.0027702362276613712
step: 80, loss: 0.0018135200953111053
step: 90, loss: 0.014924442395567894
step: 100, loss: 0.003896316746249795
step: 110, loss: 0.039911653846502304
step: 120, loss: 0.006309024058282375
step: 130, loss: 0.0005584997124969959
step: 140, loss: 0.004906550049781799
step: 150, loss: 0.0009866608306765556
step: 160, loss: 0.0064641996286809444
step: 170, loss: 0.022191127762198448
step: 180, loss: 0.0028695864602923393
step: 190, loss: 0.0029537170194089413
step: 200, loss: 0.0009360265685245395
step: 210, loss: 0.0047574639320373535
step: 220, loss: 0.15624837577342987
step: 230, loss: 0.026129987090826035
/usr/local/lib/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=0.9832026875699889, f1=0.9752252252252253,
best_f1=0.9752252252252253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003732482437044382
step: 10, loss: 0.0008310648845508695
step: 20, loss: 0.04446747899055481
step: 30, loss: 0.0006045193877071142
step: 40, loss: 0.0194082111120224
step: 50, loss: 0.009870955720543861
step: 60, loss: 0.002966222818940878
step: 70, loss: 0.007349279243499041
step: 80, loss: 0.18024897575378418
step: 90, loss: 0.11105216294527054
step: 100, loss: 0.020313534885644913
step: 110, loss: 0.0025573328603059053
step: 120, loss: 0.04886141046881676
step: 130, loss: 0.05053423345088959
step: 140, loss: 0.0013563156826421618
step: 150, loss: 0.0004846763040404767
step: 160, loss: 0.09694766998291016
step: 170, loss: 0.0027918322011828423
step: 180, loss: 0.0059484136290848255
step: 190, loss: 0.008244059979915619
step: 200, loss: 0.0014625107869505882
step: 210, loss: 0.00019786509801633656
step: 220, loss: 0.0856567919254303
step: 230, loss: 0.00037276800139807165
/usr/local/lib/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=0.9842342342342343, f1=0.9751693002257337,
best_f1=0.9751693002257337
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00031291169580072165
step: 10, loss: 0.0003256092604715377
step: 20, loss: 0.00020909200247842818
step: 30, loss: 0.00025469440151937306
step: 40, loss: 0.0005060299299657345
step: 50, loss: 0.00025238405214622617
step: 60, loss: 0.0006400216370820999
step: 70, loss: 0.00019371452799532562
step: 80, loss: 0.0004325121408328414
step: 90, loss: 0.2292923778295517
step: 100, loss: 0.02959291636943817
```

```
step: 110, loss: 0.0021032304503023624
step: 120, loss: 0.065297432243824
step: 130, loss: 0.09066581726074219
step: 140, loss: 0.0002429815795039758
step: 150, loss: 0.0004956448683515191
step: 160, loss: 0.007481192238628864
step: 170, loss: 0.023899786174297333
step: 180, loss: 0.002250840188935399
step: 190, loss: 0.00025152438320219517
step: 200, loss: 0.013557166792452335
step: 210, loss: 0.0002502963470760733
step: 220, loss: 0.000571072509046644
step: 230, loss: 0.0003154033038299531
/usr/local/lib/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=0.9843400447427293, f1=0.9831649831649831,
best_f1=0.9831649831649831
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011998831760138273
step: 10, loss: 0.003176250262185931
step: 20, loss: 0.0003642861847765744
step: 30, loss: 0.0006461014272645116
step: 40, loss: 0.0060978420078754425
step: 50, loss: 0.01970956288278103
step: 60, loss: 0.015193598344922066
step: 70, loss: 0.000328941474435851
step: 80, loss: 0.00031447754008695483
step: 90, loss: 0.00010615654900902882
step: 100, loss: 0.0002487940655555576
step: 110, loss: 0.0021672323346138
step: 120, loss: 0.00018137558072339743
step: 130, loss: 0.017192302271723747
step: 140, loss: 0.0014338284963741899
step: 150, loss: 0.00030560712912119925
step: 160, loss: 0.11485396325588226
step: 170, loss: 0.00597474817186594
step: 180, loss: 0.0013932106085121632
step: 190, loss: 0.018597031012177467
step: 200, loss: 0.00029153074137866497
step: 210, loss: 0.000439291208749637
step: 220, loss: 0.00022266268206294626
```

```
step: 230, loss: 0.0015467905905097723
/usr/local/lib/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=0.9854423292273236, f1=0.9775784753363228,
best_f1=0.9775784753363228
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06554988026618958
step: 10, loss: 0.00042588220094433061
step: 20, loss: 0.00021712009038310498
step: 30, loss: 0.0005168318748474121
step: 40, loss: 0.0005945393932051957
step: 50, loss: 0.0001924028474604711
step: 60, loss: 0.0014923607232049108
step: 70, loss: 0.0002418918302282691
step: 80, loss: 8.844855619827285e-05
step: 90, loss: 0.00020602485164999962
step: 100, loss: 0.00028847536304965615
step: 110, loss: 0.0008749673725105822
step: 120, loss: 0.00038259767461568117
step: 130, loss: 0.0013319721911102533
step: 140, loss: 0.00010803597979247570
step: 150, loss: 0.00013999165093991905
step: 160, loss: 0.00010792647663038224
step: 170, loss: 7.940678187878802e-05
step: 180, loss: 0.0041318279691040520
step: 190, loss: 0.00010896709864027798
step: 200, loss: 0.00012454460375010967
step: 210, loss: 0.0011854922631755471
step: 220, loss: 0.00010034660226665437
step: 230, loss: 0.012855983339250088
/usr/local/lib/python3.7/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=0.9832026875699889, f1=0.9785794813979707,
best_f1=0.9775784753363228
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002837179694324732
step: 10, loss: 0.00036385960993357003
step: 20, loss: 0.0001754763361532241
```

```
step: 30, loss: 0.0008379104547202587
step: 40, loss: 9.19229569262825e-05
step: 50, loss: 0.000167235397384502
step: 60, loss: 0.00011958504182985052
step: 70, loss: 0.00014073778584133834
step: 80, loss: 0.00011182817979715765
step: 90, loss: 0.00010029698751168326
step: 100, loss: 0.00010573607141850516
step: 110, loss: 0.00013278886035550386
step: 120, loss: 0.0023817650508135557
step: 130, loss: 0.00068114016903356433
step: 140, loss: 0.00013922469224780798
step: 150, loss: 0.00013424435746856034
step: 160, loss: 0.000152059001266025
step: 170, loss: 0.00013498833868652582
step: 180, loss: 0.0002854917256627232
step: 190, loss: 0.0003849549393635243
step: 200, loss: 0.00033034273656085134
step: 210, loss: 0.0001500892249168828
step: 220, loss: 0.0020438097417354584
step: 230, loss: 0.013611389324069023
/usr/local/lib/python3.7/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=0.9819413092550789, f1=0.975, best_f1=0.9775784753363228
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012446653563529253
step: 10, loss: 0.00019078530021943152
step: 20, loss: 0.0002467361628077924
step: 30, loss: 0.00025151006411761045
step: 40, loss: 0.0003465350600890815
step: 50, loss: 0.0007539461948908865
step: 60, loss: 0.0001306971098529175
step: 70, loss: 0.00016092431906145066
step: 80, loss: 0.0006713948678225279
step: 90, loss: 0.00043149973498657346
step: 100, loss: 0.00012228937703184783
step: 110, loss: 0.00024125412164721638
step: 120, loss: 0.026095952838659286
step: 130, loss: 0.0008091985364444554
step: 140, loss: 0.09696516394615173
step: 150, loss: 0.00011376906331861392
step: 160, loss: 0.0008639748557470739
step: 170, loss: 0.00011348642874509096
step: 180, loss: 0.0008821541559882462
step: 190, loss: 8.263139170594513e-05
```

```
step: 200, loss: 9.341852273792028e-05
step: 210, loss: 9.581955237081274e-05
step: 220, loss: 0.014080763794481754
step: 230, loss: 0.0002711386187002063
/usr/local/lib/python3.7/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=0.9818594104308391, f1=0.9736540664375716,
best_f1=0.9775784753363228
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011692530824802816
step: 10, loss: 0.00016414029232691973
step: 20, loss: 0.0008041763794608414
step: 30, loss: 0.00012776443327311426
step: 40, loss: 6.544789357576519e-05
step: 50, loss: 0.00019533029990270734
step: 60, loss: 0.0036453562788665295
step: 70, loss: 0.0001457364414818585
step: 80, loss: 0.00010539311915636063
step: 90, loss: 0.00015528833318967372
step: 100, loss: 8.396641351282597e-05
step: 110, loss: 0.00013340440636966377
step: 120, loss: 6.648981798207387e-05
step: 130, loss: 0.018583334982395172
step: 140, loss: 0.009047197178006172
step: 150, loss: 7.807001384207979e-05
step: 160, loss: 6.79419026710093e-05
step: 170, loss: 6.789436156395823e-05
step: 180, loss: 7.521337101934478e-05
step: 190, loss: 7.898605690570548e-05
step: 200, loss: 8.407295536017045e-05
step: 210, loss: 0.0001339982118224725
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 220, loss: 0.00039138595457188785
step: 230, loss: 0.000488699646666646
/usr/local/lib/python3.7/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=0.9854096520763187, f1=0.9764309764309763,
best_f1=0.9775784753363228
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012553378473967314
step: 10, loss: 0.0005745191592723131
step: 20, loss: 5.520516060641967e-05
```

```
step: 30, loss: 0.003966122400015593
step: 40, loss: 0.00011754537263186648
step: 50, loss: 0.00018868569168262184
step: 60, loss: 0.00018648353579919785
step: 70, loss: 9.464948379900306e-05
step: 80, loss: 0.06915564090013504
step: 90, loss: 0.0007266694447025657
step: 100, loss: 0.00010250193736283109
step: 110, loss: 8.691985567566007e-05
step: 120, loss: 0.001581943710334599
step: 130, loss: 8.219940355047584e-05
step: 140, loss: 0.00013624210259877145
step: 150, loss: 5.1428094593575224e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 160, loss: 0.00021633942378684878
step: 170, loss: 0.00012552959378808737
step: 180, loss: 0.00010663462308002636
step: 190, loss: 0.0003091036051046103
step: 200, loss: 9.310164750786498e-05
step: 210, loss: 5.729049371439032e-05
step: 220, loss: 0.0003216514887753874
step: 230, loss: 0.00020146180759184062
/usr/local/lib/python3.7/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=0.9808342728297633, f1=0.9773755656108598,
best_f1=0.9775784753363228
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014759012265130877
step: 10, loss: 0.00010815200221259147
step: 20, loss: 7.189728785306215e-05
step: 30, loss: 8.300054469145834e-05
step: 40, loss: 3.643567470135167e-05
step: 50, loss: 0.0002657411096151918
step: 60, loss: 4.510751386987977e-05
step: 70, loss: 0.00017056058277375996
step: 80, loss: 5.1609538786578923e-05
step: 90, loss: 0.0001074926694855094
step: 100, loss: 3.896428825100884e-05
step: 110, loss: 0.0003502533654682338
step: 120, loss: 0.0006484034820459783
step: 130, loss: 9.516992577118799e-05
step: 140, loss: 5.8292636822443455e-05
step: 150, loss: 5.2052255341550335e-05
step: 160, loss: 0.0005570926587097347
step: 170, loss: 5.5599899496883154e-05
```

```
step: 180, loss: 5.3430809202836826e-05
step: 190, loss: 0.0003281111130490899
step: 200, loss: 7.701556023675948e-05
step: 210, loss: 6.556051084771752e-05
step: 220, loss: 0.0027864710427820683
step: 230, loss: 0.0002856724022421986
/usr/local/lib/python3.7/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=0.9853768278965129, f1=0.9730941704035874,
best_f1=0.9775784753363228
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00022227955923881382
step: 10, loss: 0.00029162157443352044
step: 20, loss: 8.93172764335759e-05
step: 30, loss: 0.0011570241767913103
step: 40, loss: 0.0013600568054243922
step: 50, loss: 6.326959555735812e-05
step: 60, loss: 9.443280578125268e-05
step: 70, loss: 8.468180021736771e-05
step: 80, loss: 7.389383972622454e-05
step: 90, loss: 6.179241609061137e-05
step: 100, loss: 7.253784860949963e-05
step: 110, loss: 5.806417902931571e-05
step: 120, loss: 0.00014465878484770656
step: 130, loss: 9.996538574341685e-05
step: 140, loss: 0.00017345014202874154
step: 150, loss: 0.00011454772175056866
step: 160, loss: 3.958333400078118e-05
step: 170, loss: 5.0106154958484694e-05
step: 180, loss: 9.858186967903748e-05
step: 190, loss: 3.976199513999745e-05
step: 200, loss: 0.00012737218639813364
step: 210, loss: 0.00014135029050521553
step: 220, loss: 0.00033450551563873887
step: 230, loss: 0.0001439031766494736
/usr/local/lib/python3.7/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=0.9853768278965129, f1=0.9741863075196409,
best_f1=0.9775784753363228
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000351229275111109
step: 10, loss: 4.2005831346614286e-05
```

```
step: 20, loss: 0.00018172760610468686
step: 30, loss: 9.401877468917519e-05
step: 40, loss: 3.014779940713197e-05
step: 50, loss: 0.0042046005837619305
step: 60, loss: 3.355625813128427e-05
step: 70, loss: 6.08132359047886e-05
step: 80, loss: 8.125878230202943e-05
step: 90, loss: 0.0001272999943466857
step: 100, loss: 0.0009723974508233368
step: 110, loss: 0.0001830776600399986
step: 120, loss: 8.185773913282901e-05
step: 130, loss: 9.440123540116474e-05
step: 140, loss: 0.0003639844071585685
step: 150, loss: 0.00010982847743434832
step: 160, loss: 3.755615762202069e-05
step: 170, loss: 0.0002817735949065536
step: 180, loss: 0.00019450721447356045
step: 190, loss: 9.277762001147494e-05
step: 200, loss: 3.8540223613381386e-05
step: 210, loss: 0.025585344061255455
step: 220, loss: 7.341831224039197e-05
step: 230, loss: 3.374598964001052e-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=0.9853768278965129, f1=0.9741282339707535,
best_f1=0.9775784753363228
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.896832640748471e-05
step: 10, loss: 8.794313907856122e-05
step: 20, loss: 8.979291305877268e-05
step: 30, loss: 0.00018478416313882917
step: 40, loss: 6.75254050293006e-05
step: 50, loss: 6.442425365094095e-05
step: 60, loss: 5.2782081183977425e-05
step: 70, loss: 5.6382184993708506e-05
step: 80, loss: 0.00031221829704008996
step: 90, loss: 3.237911005271599e-05
step: 100, loss: 9.345552825834602e-05
step: 110, loss: 0.00010322032903786749
step: 120, loss: 6.096845754655078e-05
step: 130, loss: 0.00017891298921313137
step: 140, loss: 6.0012287576682866e-05
step: 150, loss: 0.00017252237012144178
step: 160, loss: 4.989938679500483e-05
step: 170, loss: 5.170223448658362e-05
```

```
step: 180, loss: 6.749524618498981e-05
step: 190, loss: 0.001305840676650405
step: 200, loss: 6.196892354637384e-05
step: 210, loss: 0.00013573860633186996
step: 220, loss: 6.207945989444852e-05
step: 230, loss: 0.0005283237551338971
/usr/local/lib/python3.7/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=0.9853768278965129, f1=0.9741282339707535,
best_f1=0.9775784753363228
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 [00:10, 232.47it/s]
load_f1 = 0.9854423292273236
real_f1 = 0.9854096520763187
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, 253.87it/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 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: 0.8113762140274048
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4784844219684601
step: 20, loss: 0.5154578685760498
step: 30, loss: 0.4702501893043518
step: 40, loss: 0.4448302388191223
step: 50, loss: 0.29702359437942505
step: 60, loss: 0.21470871567726135
step: 70, loss: 0.11468992382287979
step: 80, loss: 0.4065744876861572
step: 90, loss: 0.2901664972305298
step: 100, loss: 0.2566194236278534
step: 110, loss: 0.22361594438552856
step: 120, loss: 0.04820800945162773
step: 130, loss: 0.025366000831127167
step: 140, loss: 0.31401658058166504
step: 150, loss: 0.051234401762485504
step: 160, loss: 0.12270424515008926
step: 170, loss: 0.2862277328968048
step: 180, loss: 0.08797804266214371
step: 190, loss: 0.01298640388995409
step: 200, loss: 0.13080808520317078
step: 210, loss: 0.1289507895708084
step: 220, loss: 0.18324004113674164
step: 230, loss: 0.12075388431549072
step: 240, loss: 0.09415344148874283
step: 250, loss: 0.07287229597568512
step: 260, loss: 0.04779432341456413
```

```
step: 270, loss: 0.030978973954916
step: 280, loss: 0.11748608201742172
step: 290, loss: 0.09140928834676743
step: 300, loss: 0.04505634680390358
step: 310, loss: 0.05746915936470032
step: 320, loss: 0.06223113834857941
step: 330, loss: 0.04679457098245621
step: 340, loss: 0.22935202717781067
step: 350, loss: 0.03670969232916832
step: 360, loss: 0.0644993856549263
step: 370, loss: 0.1445297747850418
step: 380, loss: 0.21609850227832794
step: 390, loss: 0.1962672621011734
step: 400, loss: 0.016386263072490692
step: 410, loss: 0.0748356282711029
step: 420, loss: 0.05724892020225525
step: 430, loss: 0.06631507724523544
step: 440, loss: 0.24863669276237488
step: 450, loss: 0.010921070352196693
step: 460, loss: 0.06069058179855347
step: 470, loss: 0.36033758521080017
step: 480, loss: 0.15321208536624908
step: 490, loss: 0.0663444846868515
step: 500, loss: 0.0362828820943824
step: 510, loss: 0.14764176309108734
step: 520, loss: 0.1493835598230362
step: 530, loss: 0.1705339550971985
/usr/local/lib/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=0.9237209302325581, f1=0.9205759405480723,
best_f1=0.9205759405480723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0951521173119545
step: 10, loss: 0.15769469738006592
step: 20, loss: 0.11743658781051636
step: 30, loss: 0.07916419953107834
step: 40, loss: 0.03468076139688492
step: 50, loss: 0.06042297184467316
step: 60, loss: 0.08789342641830444
step: 70, loss: 0.1438581347465515
step: 80, loss: 0.015943612903356552
```

```
step: 90, loss: 0.07224833965301514
step: 100, loss: 0.20879632234573364
step: 110, loss: 0.01772243157029152
step: 120, loss: 0.17750033736228943
step: 130, loss: 0.07772552967071533
step: 140, loss: 0.025860263034701347
step: 150, loss: 0.012562127783894539
step: 160, loss: 0.08426413685083389
step: 170, loss: 0.1498340368270874
step: 180, loss: 0.004226949531584978
step: 190, loss: 0.06840480864048004
step: 200, loss: 0.0385945662856102
step: 210, loss: 0.06941839307546616
step: 220, loss: 0.07909894734621048
step: 230, loss: 0.028245270252227783
step: 240, loss: 0.12982536852359772
step: 250, loss: 0.05542668700218201
step: 260, loss: 0.11400729417800903
step: 270, loss: 0.20812992751598358
step: 280, loss: 0.25412940979003906
step: 290, loss: 0.1038801297545433
step: 300, loss: 0.024877473711967468
step: 310, loss: 0.0468335822224617
step: 320, loss: 0.10307228565216064
step: 330, loss: 0.07544483989477158
step: 340, loss: 0.07391104847192764
step: 350, loss: 0.08607038855552673
step: 360, loss: 0.11484401673078537
step: 370, loss: 0.05430511012673378
step: 380, loss: 0.05895994231104851
step: 390, loss: 0.03518592566251755
step: 400, loss: 0.059621814638376236
step: 410, loss: 0.0027323574759066105
step: 420, loss: 0.1205405443906784
step: 430, loss: 0.037419188767671585
step: 440, loss: 0.037745483219623566
step: 450, loss: 0.007361085619777441
step: 460, loss: 0.2389136701822281
step: 470, loss: 0.047062333673238754
step: 480, loss: 0.25180506706237793
step: 490, loss: 0.029656650498509407
step: 500, loss: 0.03172373026609421
step: 510, loss: 0.07539291679859161
step: 520, loss: 0.008089057169854641
step: 530, loss: 0.18669039011001587
/usr/local/lib/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=0.9299065420560748, f1=0.9173708920187793,
best_f1=0.9173708920187793
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04628777131438255
step: 10, loss: 0.02230992168188095
step: 20, loss: 0.13791346549987793
step: 30, loss: 0.17679594457149506
step: 40, loss: 0.012588356621563435
step: 50, loss: 0.0063318428583443165
step: 60, loss: 0.00335192633792758
step: 70, loss: 0.02620534785091877
step: 80, loss: 0.00990781094878912
step: 90, loss: 0.038940392434597015
step: 100, loss: 0.0025276902597397566
step: 110, loss: 0.04360514506697655
step: 120, loss: 0.03096488118171692
step: 130, loss: 0.04711683839559555
step: 140, loss: 0.10020694881677628
step: 150, loss: 0.028796343132853508
step: 160, loss: 0.0036160829477012157
step: 170, loss: 0.03380803018808365
step: 180, loss: 0.004087836015969515
step: 190, loss: 0.005615053698420525
step: 200, loss: 0.018209705129265785
step: 210, loss: 0.14467161893844604
step: 220, loss: 0.018948843702673912
step: 230, loss: 0.06774023920297623
step: 240, loss: 0.055654630064964294
step: 250, loss: 0.014887550845742226
step: 260, loss: 0.001598796108737588
step: 270, loss: 0.004534912761300802
step: 280, loss: 0.026228278875350952
step: 290, loss: 0.10363622009754181
step: 300, loss: 0.0846710056066513
step: 310, loss: 0.08910179138183594
step: 320, loss: 0.04627758264541626
step: 330, loss: 0.0025939163751900196
step: 340, loss: 0.016128966584801674
step: 350, loss: 0.06321603059768677
step: 360, loss: 0.0042930045165121555
step: 370, loss: 0.031085800379514694
step: 380, loss: 0.0013395303394645452

```
step: 390, loss: 0.021717390045523643
step: 400, loss: 0.03460314869880676
step: 410, loss: 0.02925661765038967
step: 420, loss: 0.061620257794857025
step: 430, loss: 0.11711905896663666
step: 440, loss: 0.006685616448521614
step: 450, loss: 0.07717175036668777
step: 460, loss: 0.07647411525249481
step: 470, loss: 0.0046465699560940266
step: 480, loss: 0.009796884842216969
step: 490, loss: 0.07591904699802399
step: 500, loss: 0.01142754964530468
step: 510, loss: 0.0034008172806352377
step: 520, loss: 0.00863865576684475
step: 530, loss: 0.005805075168609619
/usr/local/lib/python3.7/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=0.9296693060083837, f1=0.9235787511649579,
best_f1=0.9173708920187793
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005437428131699562
step: 10, loss: 0.02881254442036152
step: 20, loss: 0.010797644965350628
step: 30, loss: 0.07516717165708542
step: 40, loss: 0.01473093032836914
step: 50, loss: 0.009883091785013676
step: 60, loss: 0.005990379024296999
step: 70, loss: 0.0010706369066610932
step: 80, loss: 0.010613305494189262
step: 90, loss: 0.0004033962031826377
step: 100, loss: 0.17749786376953125
step: 110, loss: 0.0028251400217413902
step: 120, loss: 0.004273891914635897
step: 130, loss: 0.055127255618572235
step: 140, loss: 0.08700410276651382
step: 150, loss: 0.001388733508065343
step: 160, loss: 0.002034583827480674
step: 170, loss: 0.0009492789977230132
step: 180, loss: 0.006236725486814976
step: 190, loss: 0.0031258310191333294
step: 200, loss: 0.00045453631901182234
step: 210, loss: 0.17350010573863983
step: 220, loss: 0.02776309661567211
step: 230, loss: 0.28927111625671387
step: 240, loss: 0.003680844558402896
```

```
step: 250, loss: 0.008460001088678837
step: 260, loss: 0.10619201511144638
step: 270, loss: 0.05845247581601143
step: 280, loss: 0.04076629877090454
step: 290, loss: 0.11438090354204178
step: 300, loss: 0.0016319690039381385
step: 310, loss: 0.01811143383383751
step: 320, loss: 0.022980377078056335
step: 330, loss: 0.007702283561229706
step: 340, loss: 0.131557434797287
step: 350, loss: 0.04757612571120262
step: 360, loss: 0.02760600857436657
step: 370, loss: 0.03912633657455444
step: 380, loss: 0.004868061747401953
step: 390, loss: 0.1799089014530182
step: 400, loss: 0.008795109577476978
step: 410, loss: 0.0046594832092523575
step: 420, loss: 0.015223205089569092
step: 430, loss: 0.014671536162495613
step: 440, loss: 0.07440730184316635
step: 450, loss: 0.06546963006258011
step: 460, loss: 0.0208248570561409
step: 470, loss: 0.002717739436775446
step: 480, loss: 0.003174121491611004
step: 490, loss: 0.03300326690077782
step: 500, loss: 0.003854514332488179
step: 510, loss: 0.009725196287035942
step: 520, loss: 0.08399172127246857
step: 530, loss: 0.0010042733047157526
/usr/local/lib/python3.7/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=0.9272811486799445, f1=0.9159663865546218,
best_f1=0.9173708920187793
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017465690150856972
step: 10, loss: 0.054363999515771866
step: 20, loss: 0.0036913778167217797
step: 30, loss: 0.007137195207178593
step: 40, loss: 0.01219919789582491
step: 50, loss: 0.001791505143046379
step: 60, loss: 0.00054271484259516
step: 70, loss: 0.01550337765365839
step: 80, loss: 0.0013231687480583787
step: 90, loss: 0.00100766972173300534
step: 100, loss: 0.0008322977810166776
```

```
step: 110, loss: 0.030905628576874733
step: 120, loss: 0.00817389041185379
step: 130, loss: 0.003356433240696788
step: 140, loss: 0.0005901337135583162
step: 150, loss: 0.11685174703598022
step: 160, loss: 0.023913243785500526
step: 170, loss: 0.1205136626958847
step: 180, loss: 0.02849297598004341
step: 190, loss: 0.000741812284104526
step: 200, loss: 0.037063080817461014
step: 210, loss: 0.010398289188742638
step: 220, loss: 0.0011209654621779919
step: 230, loss: 0.0017115628579631448
step: 240, loss: 0.008012816309928894
step: 250, loss: 0.0001081475056707859
step: 260, loss: 0.00033896180684678257
step: 270, loss: 0.025797612965106964
step: 280, loss: 0.020294401794672012
step: 290, loss: 0.00041862670332193375
step: 300, loss: 0.006458187010139227
step: 310, loss: 0.020491711795330048
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 320, loss: 0.16213662922382355
step: 330, loss: 0.010414030402898788
step: 340, loss: 0.04465818777680397
step: 350, loss: 0.02366587333381176
step: 360, loss: 0.01656379923224449
step: 370, loss: 0.0018262909725308418
step: 380, loss: 0.0345403254032135
step: 390, loss: 0.002521356102079153
step: 400, loss: 0.02442917227745056
step: 410, loss: 0.00017558532999828458
step: 420, loss: 0.002322850516065955
step: 430, loss: 0.007089600898325443
step: 440, loss: 0.003801352344453335
step: 450, loss: 0.03987564519047737
step: 460, loss: 0.0013359605800360441
step: 470, loss: 0.0049729724414646626
step: 480, loss: 0.007010065484791994
step: 490, loss: 0.0006587684620171785
step: 500, loss: 0.00202204124070704
step: 510, loss: 0.008839140646159649
step: 520, loss: 0.000147185186506249
step: 530, loss: 0.017935194075107574
/usr/local/lib/python3.7/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=0.9197080291970803, f1=0.9157175398633257,
```

```
best_f1=0.9173708920187793
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010843687923625112
step: 10, loss: 0.0006254528998397291
step: 20, loss: 0.013514458201825619
step: 30, loss: 0.0030554733239114285
step: 40, loss: 0.00043709538294933736
step: 50, loss: 0.00015094179252628237
step: 60, loss: 8.920070831663907e-05
step: 70, loss: 0.04538016393780708
step: 80, loss: 0.00027977541321888566
step: 90, loss: 0.007333600893616676
step: 100, loss: 0.009588627144694328
step: 110, loss: 0.0020094658248126507
step: 120, loss: 0.0008256105938926339
step: 130, loss: 0.00012441715807653964
step: 140, loss: 0.0004638110694941133
step: 150, loss: 0.0073075308464467525
step: 160, loss: 8.329121192218736e-05
step: 170, loss: 0.0001309553481405601
step: 180, loss: 0.0001273617526749149
step: 190, loss: 0.0002638362639117986
step: 200, loss: 0.0006183033110573888
step: 210, loss: 0.0002651666582096368
step: 220, loss: 0.0007974799373187125
step: 230, loss: 0.00010830147220985964
step: 240, loss: 0.004841746762394905
step: 250, loss: 0.0019156449707224965
step: 260, loss: 0.0017543904250487685
step: 270, loss: 0.0007448194664902985
step: 280, loss: 0.0014826193219050765
step: 290, loss: 0.003948959056288004
step: 300, loss: 0.02676292322576046
step: 310, loss: 0.0002639865851961076
step: 320, loss: 0.004184278659522533
step: 330, loss: 0.0007263273582793772
step: 340, loss: 0.05455433949828148
step: 350, loss: 0.0007651453488506377
step: 360, loss: 0.00013153997133485973
step: 370, loss: 0.0010302551090717316
step: 380, loss: 0.010017609223723412
step: 390, loss: 0.02693796344101429
step: 400, loss: 0.0016059045447036624
step: 410, loss: 0.005470859818160534
step: 420, loss: 0.012707718648016453
step: 430, loss: 0.00045034277718514204
```

```
step: 440, loss: 0.004505612887442112
step: 450, loss: 0.006351111456751823
step: 460, loss: 0.08545956760644913
step: 470, loss: 0.03461140766739845
step: 480, loss: 0.009872877039015293
step: 490, loss: 0.001120265806093812
step: 500, loss: 0.006740370765328407
step: 510, loss: 0.001936888787895441
step: 520, loss: 0.003268341301009059
step: 530, loss: 0.0032351678237318993
/usr/local/lib/python3.7/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=0.9219467401285583, f1=0.911508482347547,
best_f1=0.9173708920187793
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0093734972178936
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 10, loss: 0.01593134179711342
step: 20, loss: 0.005553595256060362
step: 30, loss: 0.002339961240068078
step: 40, loss: 0.0072409603744745255
step: 50, loss: 0.009040839038789272
step: 60, loss: 0.008746213279664516
step: 70, loss: 0.00751599483191967
step: 80, loss: 0.0004873050784226507
step: 90, loss: 0.00015278402133844793
step: 100, loss: 0.00021778537484351546
step: 110, loss: 0.03458445519208908
step: 120, loss: 9.665856487117708e-05
step: 130, loss: 6.630923599004745e-05
step: 140, loss: 0.0007617934024892747
step: 150, loss: 7.988922880031168e-05
step: 160, loss: 0.0004067129921168089
step: 170, loss: 0.00041627982864156365
step: 180, loss: 0.0025899994652718306
step: 190, loss: 0.00011161908332724124
step: 200, loss: 0.006139947567135096
step: 210, loss: 0.0010973612079396844
step: 220, loss: 0.0004054573946632445
step: 230, loss: 0.04464361071586609
step: 240, loss: 0.006184146273881197
step: 250, loss: 0.0001943485694937408
step: 260, loss: 0.01201644167304039
step: 270, loss: 5.1462517149047926e-05
step: 280, loss: 0.0031707461457699537
```

```
step: 290, loss: 0.026900535449385643
step: 300, loss: 0.0027515795081853867
step: 310, loss: 0.0005275348084978759
step: 320, loss: 0.027546476572752
step: 330, loss: 0.0004808206285815686
step: 340, loss: 0.0037223692052066326
step: 350, loss: 0.008288071490824223
step: 360, loss: 0.0020678399596363306
step: 370, loss: 0.00010497920447960496
step: 380, loss: 0.013564216904342175
step: 390, loss: 0.00018710608128458261
step: 400, loss: 6.840204878244549e-05
step: 410, loss: 0.0006807809695601463
step: 420, loss: 0.0004759035655297339
step: 430, loss: 9.003344894153997e-05
step: 440, loss: 0.0002673541894182563
step: 450, loss: 0.0007773403776809573
step: 460, loss: 0.001525639439933002
step: 470, loss: 0.00118173286318779
step: 480, loss: 5.8256511692889035e-05
step: 490, loss: 0.00010341423330828547
step: 500, loss: 0.0001408284588251263
step: 510, loss: 0.00032227981137111783
step: 520, loss: 0.00013930359273217618
step: 530, loss: 7.040308264549822e-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=0.9296693060083837, f1=0.9186424918642492,
best_f1=0.9173708920187793
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.121568680508062e-05
step: 10, loss: 0.00037852866807952523
step: 20, loss: 0.001630857470445335
step: 30, loss: 0.00010195621871389449
step: 40, loss: 8.947131573222578e-05
step: 50, loss: 0.00017624614702071995
step: 60, loss: 0.0022349481005221605
step: 70, loss: 0.23660817742347717
step: 80, loss: 0.00033812489709816873
step: 90, loss: 0.0005933560314588249
step: 100, loss: 0.00030124178738333285
step: 110, loss: 0.0416744239628315
step: 120, loss: 0.000628973008133471
step: 130, loss: 0.0006316957296803594
step: 140, loss: 4.560931847663596e-05
```

```
step: 150, loss: 0.0008021125686354935
step: 160, loss: 0.023662051185965538
step: 170, loss: 0.00013809798110742122
step: 180, loss: 0.15999259054660797
step: 190, loss: 0.0010352766839787364
step: 200, loss: 0.008404778316617012
step: 210, loss: 0.02886199951171875
step: 220, loss: 0.001788407564163208
step: 230, loss: 0.008802266791462898
step: 240, loss: 0.0011175135150551796
step: 250, loss: 0.0011335633462294936
step: 260, loss: 0.04007073491811752
step: 270, loss: 0.000639108067844063
step: 280, loss: 0.0012099642772227526
step: 290, loss: 0.0024082395248115063
step: 300, loss: 0.0001634690270293504
step: 310, loss: 0.02951047010719776
step: 320, loss: 0.000211590071558021
step: 330, loss: 0.022595606744289398
step: 340, loss: 0.0002696568553801626
step: 350, loss: 0.00062192557379960339
step: 360, loss: 5.638176662614569e-05
step: 370, loss: 0.0016527153784409165
step: 380, loss: 0.00035746817593462765
step: 390, loss: 0.017960140481591225
step: 400, loss: 0.000618673162534833
step: 410, loss: 0.003498287871479988
step: 420, loss: 5.7953882787842304e-05
step: 430, loss: 0.2413928508758545
step: 440, loss: 0.0016967904521152377
step: 450, loss: 0.00019064194930251688
step: 460, loss: 0.003521036822348833
step: 470, loss: 0.014510050415992737
step: 480, loss: 0.0004759521980304271
step: 490, loss: 0.00011591726070037112
step: 500, loss: 0.0002380572841502726
step: 510, loss: 0.000245720031671226
step: 520, loss: 0.000128196959849447
step: 530, loss: 0.00018383015412837267
/usr/local/lib/python3.7/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=0.9265381083562901, f1=0.9205909510618652,
best_f1=0.9173708920187793
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001544651750009507
```

```
step: 10, loss: 0.0009748446173034608
step: 20, loss: 0.0002615498087834567
step: 30, loss: 0.000412807974498719
step: 40, loss: 0.0005927723832428455
step: 50, loss: 0.009719593450427055
step: 60, loss: 0.0002835536142811179
step: 70, loss: 0.11671736091375351
step: 80, loss: 0.00011471177276689559
step: 90, loss: 0.00040232238825410604
step: 100, loss: 0.00047480763169005513
step: 110, loss: 0.004157472401857376
step: 120, loss: 0.008245792239904404
step: 130, loss: 0.0007308610365726054
step: 140, loss: 0.0014769298722967505
step: 150, loss: 0.0004591541364789009
step: 160, loss: 0.001225277315825224
step: 170, loss: 8.511515625286847e-05
step: 180, loss: 0.00013267097529023886
step: 190, loss: 0.0009453052189201117
step: 200, loss: 0.00016106142720673233
step: 210, loss: 9.146841330220923e-05
step: 220, loss: 0.0002196774585172534
step: 230, loss: 7.58612368372269e-05
step: 240, loss: 0.0006134871509857476
step: 250, loss: 0.00021585074136964977
step: 260, loss: 0.00035625582677312195
step: 270, loss: 0.002057085745036602
step: 280, loss: 6.498361472040415e-05
step: 290, loss: 5.0367569201625884e-05
step: 300, loss: 0.00032556671067140996
step: 310, loss: 7.993275357875973e-05
step: 320, loss: 0.0007148546865209937
step: 330, loss: 0.00010398765152785927
step: 340, loss: 0.00012790225446224213
step: 350, loss: 5.6699329434195533e-05
step: 360, loss: 0.001921423594467342
step: 370, loss: 0.00016957882326096296
step: 380, loss: 4.9435435357736424e-05
step: 390, loss: 5.254101779428311e-05
step: 400, loss: 0.00017284249770455062
step: 410, loss: 0.0001452933938708...563
step: 420, loss: 0.0015484687173739076
step: 430, loss: 0.00011467502190498635
step: 440, loss: 0.00037469356902875006
step: 450, loss: 0.00010339982691220939
step: 460, loss: 0.01541051920503378
step: 470, loss: 0.00011403155076565032
step: 480, loss: 0.0002932904753834009
```

```
step: 490, loss: 0.007262866012752056
step: 500, loss: 0.0013127755373716354
step: 510, loss: 7.100329821696505e-05
step: 520, loss: 0.002027201699092984
step: 530, loss: 0.00014568118785973638
/usr/local/lib/python3.7/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=0.9240150093808631, f1=0.9155261915998112,
best_f1=0.9173708920187793
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006806683959439397
step: 10, loss: 0.002360219368711114
step: 20, loss: 0.0002540685818530619
step: 30, loss: 0.00018156939768232405
step: 40, loss: 0.00015039400022942573
step: 50, loss: 8.710614201845601e-05
step: 60, loss: 0.00033297704067081213
step: 70, loss: 0.0007379840826615691
step: 80, loss: 0.001500754151493311
step: 90, loss: 0.00047960475785657763
step: 100, loss: 0.0001819539611460641
step: 110, loss: 0.0010075925383716822
step: 120, loss: 0.00027514947578310966
step: 130, loss: 0.00016575490008108318
step: 140, loss: 0.000707303814124316
step: 150, loss: 0.0001419930486008525
step: 160, loss: 0.000613152573350817
step: 170, loss: 0.0006566484225913882
step: 180, loss: 0.000407547689974308
step: 190, loss: 0.00014266495418269187
step: 200, loss: 7.93799408711493e-05
step: 210, loss: 5.544692976400256e-05
step: 220, loss: 0.00012967755901627243
step: 230, loss: 0.0001251350186066702
step: 240, loss: 0.0003969369572587311
step: 250, loss: 7.729786011623219e-05
step: 260, loss: 0.024722693488001823
step: 270, loss: 7.089643622748554e-05
step: 280, loss: 0.00012765398423653096
step: 290, loss: 5.311645509209484e-05
step: 300, loss: 0.00010331239900551736
step: 310, loss: 0.00027724800747819245
step: 320, loss: 0.00016256111848633736
step: 330, loss: 8.795969188213348e-05
step: 340, loss: 0.00028966600075364113
```

```
step: 350, loss: 0.001285618869587779
step: 360, loss: 8.830565639073029e-05
step: 370, loss: 0.0012831584317609668
step: 380, loss: 9.23930638236925e-05
step: 390, loss: 5.24433926329948e-05
step: 400, loss: 9.641076758271083e-05
step: 410, loss: 0.0005300605553202331
step: 420, loss: 0.007368441671133041
step: 430, loss: 4.28984931204468e-05
step: 440, loss: 6.062736429157667e-05
step: 450, loss: 0.0024745904374867678
step: 460, loss: 0.06805356591939926
step: 470, loss: 6.49526555207558e-05
step: 480, loss: 0.006990163121372461
step: 490, loss: 0.06386560946702957
step: 500, loss: 0.0002145848557120189
step: 510, loss: 0.00023934233468025923
step: 520, loss: 0.00012590379628818482
step: 530, loss: 0.00017488564481027424
/usr/local/lib/python3.7/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=0.9255571360834518, f1=0.913064133016627,
best_f1=0.9173708920187793
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.612886500079185e-05
step: 10, loss: 0.003461678046733141
step: 20, loss: 3.718056177604012e-05
step: 30, loss: 0.0002561853325460106
step: 40, loss: 0.00015616053133271635
step: 50, loss: 0.0009128348319791257
step: 60, loss: 0.0015018931590020657
step: 70, loss: 0.00043837440898641944
step: 80, loss: 0.000591360789258033
step: 90, loss: 0.00014156923862174153
step: 100, loss: 6.294391641858965e-05
step: 110, loss: 6.108357774792239e-05
step: 120, loss: 0.00019788501958882842
step: 130, loss: 9.772279736353084e-05
step: 140, loss: 5.084215081296861e-05
step: 150, loss: 7.944247772684321e-05
step: 160, loss: 7.573082984890789e-05
step: 170, loss: 0.004104690160602331
step: 180, loss: 0.00011220968735869974
step: 190, loss: 0.0004842059570364654
step: 200, loss: 0.000988274347037077
```

```
step: 210, loss: 0.0004895113524980843
step: 220, loss: 0.00012173347931820899
step: 230, loss: 7.490247662644833e-05
step: 240, loss: 0.00013158041110727936
step: 250, loss: 0.0007080658106133342
step: 260, loss: 0.00010336235573049635
step: 270, loss: 0.0007072556763887405
step: 280, loss: 0.0002459450624883175
step: 290, loss: 0.0022804548498243093
step: 300, loss: 0.0002601444721221924
step: 310, loss: 0.01445423811674118
step: 320, loss: 0.000837273895740509
step: 330, loss: 0.0008375540492124856
step: 340, loss: 0.0004571300814859569
step: 350, loss: 0.00028679144452326
step: 360, loss: 0.00047820122563280165
step: 370, loss: 3.19504106300883e-05
step: 380, loss: 4.760632509714924e-05
step: 390, loss: 0.0075338007882237434
step: 400, loss: 5.367948688217439e-05
step: 410, loss: 7.383709453279153e-05
step: 420, loss: 0.0007161060348153114
step: 430, loss: 0.00030342675745487213
step: 440, loss: 0.0003916777204722166
step: 450, loss: 9.796197264222428e-05
step: 460, loss: 0.0009169378317892551
step: 470, loss: 0.0005732293939217925
step: 480, loss: 0.02562624029815197
step: 490, loss: 9.152897837338969e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 500, loss: 0.000853378267493099
step: 510, loss: 0.0004026989627163857
step: 520, loss: 5.517256067832932e-05
step: 530, loss: 3.797775934799574e-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 11: dev_f1=0.9311778290993071, f1=0.9160092807424594,
best_f1=0.9160092807424594
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013223891437519342
step: 10, loss: 9.098279406316578e-05
```

```
step: 20, loss: 4.849132528761402e-05
step: 30, loss: 7.938533963169903e-05
step: 40, loss: 0.00011241822357987985
step: 50, loss: 7.736470433883369e-05
step: 60, loss: 9.471365046920255e-05
step: 70, loss: 0.010228384286165237
step: 80, loss: 0.0005613715620711446
step: 90, loss: 0.010894781909883022
step: 100, loss: 0.0001526592968730256
step: 110, loss: 4.4771535613108426e-05
step: 120, loss: 6.443490565288812e-05
step: 130, loss: 8.761427307035774e-05
step: 140, loss: 6.034085890860297e-05
step: 150, loss: 0.0023183750454336405
step: 160, loss: 0.012494487687945366
step: 170, loss: 0.0002747100661508739
step: 180, loss: 0.00035765720531344414
step: 190, loss: 2.5856630600173958e-05
step: 200, loss: 4.319292929721996e-05
step: 210, loss: 0.00015750750026199967
step: 220, loss: 5.25329596712254e-05
step: 230, loss: 0.001751564210280776
step: 240, loss: 4.210436236462556e-05
step: 250, loss: 0.0001547501451568678
step: 260, loss: 0.00010492109868209809
step: 270, loss: 3.145153823425062e-05
step: 280, loss: 7.427079981425777e-05
step: 290, loss: 0.0031426148489117622
step: 300, loss: 6.169240805320442e-05
step: 310, loss: 7.068964623613283e-05
step: 320, loss: 9.946159843821079e-05
step: 330, loss: 4.3039814045187086e-05
step: 340, loss: 0.00014616751286666684
step: 350, loss: 0.0015843837754800916
step: 360, loss: 0.00022371056547854096
step: 370, loss: 0.00016433012206107378
step: 380, loss: 0.02139795757830143
step: 390, loss: 4.262616130290553e-05
step: 400, loss: 4.2972955270670354e-05
step: 410, loss: 0.0006756422226317227
step: 420, loss: 0.0027216633316129446
step: 430, loss: 0.00024548795772716403
step: 440, loss: 0.00022995361359789968
step: 450, loss: 0.001183661282993853
step: 460, loss: 0.00011143435403937474
step: 470, loss: 0.022397466003894806
step: 480, loss: 0.0004724679747596383
step: 490, loss: 3.0021359634702094e-05
```

```
step: 500, loss: 0.009940946474671364
step: 510, loss: 8.685416833031923e-05
step: 520, loss: 0.00012518797302618623
step: 530, loss: 0.00011187911877641454
/usr/local/lib/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=0.931975937066173, f1=0.9178338001867413,
best_f1=0.9178338001867413
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.752124318154529e-05
step: 10, loss: 0.00010350849333917722
step: 20, loss: 0.000696988427080214
step: 30, loss: 0.0013502826914191246
step: 40, loss: 6.298593507381156e-05
step: 50, loss: 5.374844840844162e-05
step: 60, loss: 4.861582419835031e-05
step: 70, loss: 7.417936285492033e-05
step: 80, loss: 4.080835060449317e-05
step: 90, loss: 0.0005278587341308594
step: 100, loss: 0.0002283344801981002
step: 110, loss: 0.00012623488146346062
step: 120, loss: 7.02858014847152e-05
step: 130, loss: 0.0003130868717562407
step: 140, loss: 4.31891530752182e-05
step: 150, loss: 4.24630670750048e-05
step: 160, loss: 4.9161597416969016e-05
step: 170, loss: 0.0003780967672355473
step: 180, loss: 0.04480752348899841
step: 190, loss: 0.00011709431419149041
step: 200, loss: 0.0003525123174767941
step: 210, loss: 0.00016790663357824087
step: 220, loss: 3.456205377005972e-05
step: 230, loss: 2.7067328119301237e-05
step: 240, loss: 6.153334106784314e-05
step: 250, loss: 0.0016009919345378876
step: 260, loss: 4.6061879402259365e-05
step: 270, loss: 4.852416896028444e-05
step: 280, loss: 0.0004511014558374882
step: 290, loss: 0.0016917419852688909
step: 300, loss: 5.706253432435915e-05
step: 310, loss: 0.0002500562695786357
```

```
step: 320, loss: 7.46720252209343e-05
step: 330, loss: 0.00012580945622175932
step: 340, loss: 3.523271152516827e-05
step: 350, loss: 0.0010523315286263824
step: 360, loss: 3.579115946195088e-05
step: 370, loss: 3.637454938143492e-05
step: 380, loss: 0.0011599359568208456
step: 390, loss: 4.116599666303955e-05
step: 400, loss: 5.3615767683368176e-05
step: 410, loss: 3.307905717520043e-05
step: 420, loss: 4.0272210753755644e-05
step: 430, loss: 0.00024918594863265753
step: 440, loss: 0.0001391154364682734
step: 450, loss: 5.420290835900232e-05
step: 460, loss: 3.199902857886627e-05
step: 470, loss: 0.0068199895322322845
step: 480, loss: 6.790702900616452e-05
step: 490, loss: 5.171047814656049e-05
step: 500, loss: 6.486717757070437e-05
step: 510, loss: 0.0001126567367464304
step: 520, loss: 0.00012087567301932722
step: 530, loss: 0.00011275713768554851
/usr/local/lib/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=0.9335192933519293, f1=0.9176029962546816,
best_f1=0.9176029962546816
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006647574482485652
step: 10, loss: 8.375335164600983e-05
step: 20, loss: 3.373506842763163e-05
step: 30, loss: 6.325429421849549e-05
step: 40, loss: 6.523545016534626e-05
step: 50, loss: 0.0003914797562174499
step: 60, loss: 0.0017361952923238277
step: 70, loss: 4.214171349303797e-05
step: 80, loss: 5.429540033219382e-05
step: 90, loss: 2.651567046996206e-05
step: 100, loss: 5.317102113622241e-05
step: 110, loss: 3.17048434226308e-05
step: 120, loss: 3.350944461999461e-05
step: 130, loss: 5.318424155120738e-05
```

```
step: 140, loss: 0.026927102357149124
step: 150, loss: 0.00011849572183564305
step: 160, loss: 9.257396595785394e-05
step: 170, loss: 5.33554884896148e-05
step: 180, loss: 0.00025303979055024683
step: 190, loss: 5.3879120969213545e-05
step: 200, loss: 6.185520032886416e-05
step: 210, loss: 8.857934881234542e-05
step: 220, loss: 0.0001966080308193341
step: 230, loss: 0.0005814391770400107
step: 240, loss: 0.00036879867548123
step: 250, loss: 4.713234375230968e-05
step: 260, loss: 2.726087586779613e-05
step: 270, loss: 0.0002552050573285669
step: 280, loss: 3.6900695704389364e-05
step: 290, loss: 5.783783853985369e-05
step: 300, loss: 2.7138077712152153e-05
step: 310, loss: 0.00016490850248374045
step: 320, loss: 0.0037090752739459276
step: 330, loss: 6.280489469645545e-05
step: 340, loss: 0.0021667450200766325
step: 350, loss: 8.494422218063846e-05
step: 360, loss: 0.011028429493308067
step: 370, loss: 4.2385654523968697e-05
step: 380, loss: 3.354443470016122e-05
step: 390, loss: 4.438830728759058e-05
step: 400, loss: 0.0004228698380757123
step: 410, loss: 3.488587026367895e-05
step: 420, loss: 3.5331850085640326e-05
step: 430, loss: 9.610325651010498e-05
step: 440, loss: 3.5805103834718466e-05
step: 450, loss: 3.168983312207274e-05
step: 460, loss: 3.553790884325281e-05
step: 470, loss: 3.669964644359425e-05
step: 480, loss: 2.7931433578487486e-05
step: 490, loss: 0.0001873468136182055
step: 500, loss: 6.921339081600308e-05
step: 510, loss: 3.2445761462440714e-05
step: 520, loss: 0.00021560124878305942
step: 530, loss: 3.0594990676036105e-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=0.9225653206650831, f1=0.912280701754386,
best_f1=0.9176029962546816
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 6.378136458806694e-05
step: 10, loss: 0.00013138579379301518
step: 20, loss: 4.037089092889801e-05
step: 30, loss: 6.849748751847073e-05
step: 40, loss: 2.4355162167921662e-05
step: 50, loss: 9.011306246975437e-05
step: 60, loss: 6.26615947112441e-05
step: 70, loss: 4.6685403503943235e-05
step: 80, loss: 7.448393444065005e-05
step: 90, loss: 9.428821067558601e-05
step: 100, loss: 4.691424692282453e-05
step: 110, loss: 0.0012020396534353495
step: 120, loss: 0.001162433298304677
step: 130, loss: 2.495843727956526e-05
step: 140, loss: 2.2280330085777678e-05
step: 150, loss: 0.00011947249004151672
step: 160, loss: 3.450569784035906e-05
step: 170, loss: 4.540435111266561e-05
step: 180, loss: 4.150126551394351e-05
step: 190, loss: 7.175313658080995e-05
step: 200, loss: 6.788255268475041e-05
step: 210, loss: 4.078407437191345e-05
step: 220, loss: 2.3759088435326703e-05
step: 230, loss: 3.2345324143534526e-05
step: 240, loss: 0.00021853313955944031
step: 250, loss: 3.704210394062102e-05
step: 260, loss: 5.2094725106144324e-05
step: 270, loss: 9.810503252083436e-05
step: 280, loss: 7.181398541433737e-05
step: 290, loss: 0.016436541453003883
step: 300, loss: 0.011957943439483643
step: 310, loss: 0.00016085598326753825
step: 320, loss: 4.107614222448319e-05
step: 330, loss: 0.0019487416138872504
step: 340, loss: 4.0174953028326854e-05
step: 350, loss: 4.023115980089642e-05
step: 360, loss: 0.0001405587390763685
step: 370, loss: 3.067632133024745e-05
step: 380, loss: 3.2442108931718394e-05
step: 390, loss: 6.348300667013973e-05
step: 400, loss: 4.678760524257086e-05
step: 410, loss: 8.052628254517913e-05
step: 420, loss: 6.3491657783743e-05
step: 430, loss: 3.0781186069361866e-05
step: 440, loss: 5.3547482821159065e-05
step: 450, loss: 0.00010302884038537741
step: 460, loss: 2.5078008548007347e-05
step: 470, loss: 7.27627266314812e-05
```

```
step: 480, loss: 2.4638529794174246e-05
step: 490, loss: 3.365627708262764e-05
step: 500, loss: 6.671297887805849e-05
step: 510, loss: 0.00016337999841198325
step: 520, loss: 2.8948406907147728e-05
step: 530, loss: 4.3497006117831916e-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=0.9316081330868762, f1=0.9162790697674419,
best_f1=0.9176029962546816
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.")
5742it [00:22, 254.83it/s]
load_f1 = 0.9369453526389537
real_f1 = 0.9325842696629213
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, 257.93it/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 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: 0.8327840566635132
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.06237488240003586
step: 20, loss: 0.3870389461517334
step: 30, loss: 0.37523964047431946
step: 40, loss: 0.5109742879867554
step: 50, loss: 0.2990797758102417
step: 60, loss: 0.37046048045158386
step: 70, loss: 0.2569812536239624
step: 80, loss: 0.3025326728820801
step: 90, loss: 0.41892552375793457
step: 100, loss: 0.17302219569683075
step: 110, loss: 0.32420477271080017
step: 120, loss: 0.24942180514335632
step: 130, loss: 0.2575520873069763
step: 140, loss: 0.2852690517902374
step: 150, loss: 0.28559571504592896
step: 160, loss: 0.2366052269935608
step: 170, loss: 0.1779283583164215
step: 180, loss: 0.18046049773693085
step: 190, loss: 0.22297467291355133
step: 200, loss: 0.18063171207904816
step: 210, loss: 0.481444388628006
/usr/local/lib/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=0.488245931283906, f1=0.48314606741573035,
best_f1=0.48314606741573035
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10037596523761749
step: 10, loss: 0.0879872739315033
step: 20, loss: 0.20329825580120087
step: 30, loss: 0.18978236615657806
step: 40, loss: 0.05804048851132393
step: 50, loss: 0.2521280348300934
step: 60, loss: 0.08920931816101074
step: 70, loss: 0.22389203310012817
step: 80, loss: 0.2254820168018341
step: 90, loss: 0.1438141018152237
step: 100, loss: 0.12944892048835754
step: 110, loss: 0.08658159524202347
step: 120, loss: 0.16811393201351166
step: 130, loss: 0.2474135309457779
step: 140, loss: 0.26958605647087097
step: 150, loss: 0.2390945702791214
step: 160, loss: 0.1498931646347046
step: 170, loss: 0.25615060329437256
step: 180, loss: 0.2435726523399353
step: 190, loss: 0.2204698920249939
step: 200, loss: 0.136483296751976
step: 210, loss: 0.24914959073066711
/usr/local/lib/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=0.5548902195608783, f1=0.5489361702127661,
best_f1=0.5489361702127661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31377631425857544
step: 10, loss: 0.20552442967891693
step: 20, loss: 0.24669891595840454
step: 30, loss: 0.051035698503255844
step: 40, loss: 0.22794245183467865
step: 50, loss: 0.17261600494384766
step: 60, loss: 0.16535130143165588
step: 70, loss: 0.09629662334918976
```

```
step: 80, loss: 0.13321569561958313
step: 90, loss: 0.1564415544271469
step: 100, loss: 0.04123498126864433
step: 110, loss: 0.14171306788921356
step: 120, loss: 0.05887402221560478
step: 130, loss: 0.22925350069999695
step: 140, loss: 0.26105013489723206
step: 150, loss: 0.2129449099302292
step: 160, loss: 0.11344295740127563
step: 170, loss: 0.21341753005981445
step: 180, loss: 0.07137523591518402
step: 190, loss: 0.447238564491272
step: 200, loss: 0.09990160167217255
step: 210, loss: 0.2529895007610321
/usr/local/lib/python3.7/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=0.5373134328358209, f1=0.5330661322645291,
best_f1=0.5489361702127661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2109982967376709
step: 10, loss: 0.07917363196611404
step: 20, loss: 0.10977083444595337
step: 30, loss: 0.09263912588357925
step: 40, loss: 0.09561826288700104
step: 50, loss: 0.11310409754514694
step: 60, loss: 0.03826262056827545
step: 70, loss: 0.20032203197479248
step: 80, loss: 0.13386115431785583
step: 90, loss: 0.04281548038125038
step: 100, loss: 0.12248028814792633
step: 110, loss: 0.19681291282176971
step: 120, loss: 0.09003028273582458
step: 130, loss: 0.19327425956726074
step: 140, loss: 0.11831741780042648
step: 150, loss: 0.17534664273262024
step: 160, loss: 0.030621830374002457
step: 170, loss: 0.027972273528575897
step: 180, loss: 0.17078371345996857
step: 190, loss: 0.13628455996513367
step: 200, loss: 0.08953144401311874
step: 210, loss: 0.03170153498649597
/usr/local/lib/python3.7/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=0.5370741482965931, f1=0.537117903930131,
```

```
best_f1=0.5489361702127661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0719900131225586
step: 10, loss: 0.0474645160138607
step: 20, loss: 0.012511733919382095
step: 30, loss: 0.04449322819709778
step: 40, loss: 0.06709571182727814
step: 50, loss: 0.12732286751270294
step: 60, loss: 0.016391023993492126
step: 70, loss: 0.07590702921152115
step: 80, loss: 0.026208871975541115
step: 90, loss: 0.047401312738657
step: 100, loss: 0.13716474175453186
step: 110, loss: 0.008661805652081966
step: 120, loss: 0.09626990556716919
step: 130, loss: 0.05760183930397034
step: 140, loss: 0.020961454138159752
step: 150, loss: 0.06124265491962433
step: 160, loss: 0.08584185689687729
step: 170, loss: 0.08689907938241959
step: 180, loss: 0.09565827995538712
step: 190, loss: 0.06261187046766281
step: 200, loss: 0.1300860345363617
step: 210, loss: 0.033101391047239304
/usr/local/lib/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=0.5632798573975044, f1=0.5406427221172023,
best_f1=0.5406427221172023
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03653290495276451
step: 10, loss: 0.012064650654792786
step: 20, loss: 0.016816729679703712
step: 30, loss: 0.04268878698348999
step: 40, loss: 0.014801209792494774
step: 50, loss: 0.10106100142002106
step: 60, loss: 0.16327108442783356
step: 70, loss: 0.010018984787166119
step: 80, loss: 0.08754831552505493
step: 90, loss: 0.10563573241233826
```

```
step: 100, loss: 0.07019905745983124
step: 110, loss: 0.011270876973867416
step: 120, loss: 0.019026555120944977
step: 130, loss: 0.0110388258472085
step: 140, loss: 0.06244877725839615
step: 150, loss: 0.03190331533551216
step: 160, loss: 0.23643481731414795
step: 170, loss: 0.02434603124856949
step: 180, loss: 0.018546463921666145
step: 190, loss: 0.12203194200992584
step: 200, loss: 0.010719629935920238
step: 210, loss: 0.04425691068172455
/usr/local/lib/python3.7/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=0.5288135593220339, f1=0.5408348457350273,
best_f1=0.5406427221172023
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01070735976099968
step: 10, loss: 0.031082546338438988
step: 20, loss: 0.13361269235610962
step: 30, loss: 0.14378882944583893
step: 40, loss: 0.11514873057603836
step: 50, loss: 0.018236419185996056
step: 60, loss: 0.08145872503519058
step: 70, loss: 0.00811270996928215
step: 80, loss: 0.008289228193461895
step: 90, loss: 0.09125001728534698
step: 100, loss: 0.018431061878800392
step: 110, loss: 0.06283379346132278
step: 120, loss: 0.07142949849367142
step: 130, loss: 0.003100433386862278
step: 140, loss: 0.00719723105430603
step: 150, loss: 0.039141155779361725
step: 160, loss: 0.03415795415639877
step: 170, loss: 0.024849120527505875
step: 180, loss: 0.0014165259199216962
step: 190, loss: 0.008042898029088974
step: 200, loss: 0.030881453305482864
step: 210, loss: 0.04801696166396141
/usr/local/lib/python3.7/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=0.5423728813559322, f1=0.5011286681715574,
best_f1=0.5406427221172023
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003091208403930068
step: 10, loss: 0.036522310227155685
step: 20, loss: 0.055211860686540604
step: 30, loss: 0.005707520991563797
step: 40, loss: 0.02114705927670002
step: 50, loss: 0.006410703528672457
step: 60, loss: 0.12132573872804642
step: 70, loss: 0.0046153138391673565
step: 80, loss: 0.012913394719362259
step: 90, loss: 0.05917373672127724
step: 100, loss: 0.007611336652189493
step: 110, loss: 0.0017349474364891648
step: 120, loss: 0.006951181683689356
step: 130, loss: 0.016951339319348335
step: 140, loss: 0.0022217235527932644
step: 150, loss: 0.0496278740465641
step: 160, loss: 0.01533224806189537
step: 170, loss: 0.013868063688278198
step: 180, loss: 0.018480543047189713
step: 190, loss: 0.01955139823257923
step: 200, loss: 0.14644327759742737
step: 210, loss: 0.020616786554455757
/usr/local/lib/python3.7/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=0.5220458553791887, f1=0.5333333333333333,
best_f1=0.5406427221172023
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09596037119626999
step: 10, loss: 0.02027297578752041
step: 20, loss: 0.03886047378182411
step: 30, loss: 0.0947711244225502
step: 40, loss: 0.016022365540266037
step: 50, loss: 0.002998054726049304
step: 60, loss: 0.019480910152196884
step: 70, loss: 0.062056370079517365
step: 80, loss: 0.0025207982398569584
step: 90, loss: 0.02285122126340866
step: 100, loss: 0.0024734949693083763
step: 110, loss: 0.002136928727850318
step: 120, loss: 0.003860990284010768
step: 130, loss: 0.16901734471321106
step: 140, loss: 0.019246140494942665
step: 150, loss: 0.002595163881778717

```
step: 160, loss: 0.004905272275209427
step: 170, loss: 0.29925674200057983
step: 180, loss: 0.004459321033209562
step: 190, loss: 0.01249017659574747
step: 200, loss: 0.026989875361323357
step: 210, loss: 0.017322540283203125
/usr/local/lib/python3.7/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=0.5361216730038023, f1=0.5158562367864694,
best_f1=0.5406427221172023
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007852811366319656
step: 10, loss: 0.00367018417455256
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 20, loss: 0.017993956804275513
step: 30, loss: 0.0938861295580864
step: 40, loss: 0.024910837411880493
step: 50, loss: 0.0003042292082682252
step: 60, loss: 0.0002753391454461962
step: 70, loss: 0.010596705600619316
step: 80, loss: 0.00011524858564371243
step: 90, loss: 0.006929314229637384
step: 100, loss: 0.00021120262681506574
step: 110, loss: 0.00018699714564718306
step: 120, loss: 0.001953680766746402
step: 130, loss: 0.00068791920784860085
step: 140, loss: 0.0004563633701764047
step: 150, loss: 0.09384193271398544
step: 160, loss: 0.00030362093821167946
step: 170, loss: 0.0071649933233857155
step: 180, loss: 0.0010333918035030365
step: 190, loss: 0.027291009202599525
step: 200, loss: 0.024344954639673233
step: 210, loss: 0.02888258546590805
/usr/local/lib/python3.7/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=0.5248508946322068, f1=0.516949152542373,
best_f1=0.5406427221172023
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0024002217687666416
step: 10, loss: 0.005412280093878508
step: 20, loss: 0.0003620823554228991
```

```
step: 30, loss: 0.011202878318727016
step: 40, loss: 0.050487272441387177
step: 50, loss: 0.00024300941731780767
step: 60, loss: 0.001820664037950337
step: 70, loss: 0.000664820778183639
step: 80, loss: 0.004748060833662748
step: 90, loss: 0.06024815887212753
step: 100, loss: 0.012936031445860863
step: 110, loss: 0.0031586578115820885
step: 120, loss: 0.0003874405811075121
step: 130, loss: 0.006248250138014555
step: 140, loss: 0.08276922255754471
step: 150, loss: 0.028115293011069298
step: 160, loss: 0.002631047973409295
step: 170, loss: 0.11453711241483688
step: 180, loss: 0.0012552259722724557
step: 190, loss: 0.0010194712085649371
step: 200, loss: 0.00018630732665769756
step: 210, loss: 0.0002663304330781102
/usr/local/lib/python3.7/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=0.5414141414141413, f1=0.5258620689655173,
best_f1=0.5406427221172023
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017150533676613122
step: 10, loss: 0.000985602498985827
step: 20, loss: 0.012826531194150448
step: 30, loss: 0.010870695114135742
step: 40, loss: 0.0037385544274002314
step: 50, loss: 0.00045079554547555745
step: 60, loss: 0.010393803007900715
step: 70, loss: 0.0002296064922120422
step: 80, loss: 0.006875489838421345
step: 90, loss: 0.00012310782040003687
step: 100, loss: 0.00026453437749296427
step: 110, loss: 0.00044152920600026846
step: 120, loss: 0.0005760291242040694
step: 130, loss: 0.004045081790536642
step: 140, loss: 0.000697662471793592
step: 150, loss: 0.00023469024745281786
step: 160, loss: 0.006116851232945919
step: 170, loss: 0.002506766002625227
step: 180, loss: 0.00032991208718158305
step: 190, loss: 0.07512576133012772
step: 200, loss: 0.0020753517746925354
```

```
step: 210, loss: 0.07187389582395554
/usr/local/lib/python3.7/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=0.5294117647058824, f1=0.4867469879518072,
best_f1=0.5406427221172023
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017548519826959819
step: 10, loss: 0.00069500482641116049
step: 20, loss: 0.0020218058489263058
step: 30, loss: 0.0003241692902520299
step: 40, loss: 0.000595994177274406
step: 50, loss: 0.00016538260388188064
step: 60, loss: 0.020836755633354187
step: 70, loss: 0.026246346533298492
step: 80, loss: 0.016978060826659203
step: 90, loss: 0.001155668287537992
step: 100, loss: 0.19100145995616913
step: 110, loss: 0.00026508510927669704
step: 120, loss: 0.00036524346796795726
step: 130, loss: 0.009071502834558487
step: 140, loss: 0.0004382900951895863
step: 150, loss: 0.00026374857407063246
step: 160, loss: 0.00099537696223706
step: 170, loss: 0.08382118493318558
step: 180, loss: 0.005659355781972408
step: 190, loss: 0.00027665370726026595
step: 200, loss: 0.0004852571291849017
step: 210, loss: 0.0014340425841510296
/usr/local/lib/python3.7/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=0.5253456221198157, f1=0.46534653465346537,
best_f1=0.5406427221172023
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0025198713410645723
step: 10, loss: 0.00015973811969161034
step: 20, loss: 0.003897972870618105
step: 30, loss: 0.00016116666665766388
step: 40, loss: 0.001627512276172638
step: 50, loss: 0.021091777831315994
step: 60, loss: 0.003416430903598666
step: 70, loss: 0.0010766451014205813
step: 80, loss: 0.0009270834270864725
```

```
step: 90, loss: 0.001008115941658616
step: 100, loss: 0.0006190295680426061
step: 110, loss: 0.0004958579083904624
step: 120, loss: 0.0004994663177058101
step: 130, loss: 0.005967895500361919
step: 140, loss: 0.0007093331660144031
step: 150, loss: 0.0038702136371284723
step: 160, loss: 0.0003120255714748055
step: 170, loss: 0.04584020376205444
step: 180, loss: 0.0030717498157173395
step: 190, loss: 0.00026194073143415153
step: 200, loss: 0.00018721443484537303
step: 210, loss: 0.038135286420583725
/usr/local/lib/python3.7/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=0.5352112676056338, f1=0.47058823529411764,
best_f1=0.5406427221172023
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0047713699750602245
step: 10, loss: 0.004993115086108446
step: 20, loss: 0.0009918839205056429
step: 30, loss: 0.0007011109846644104
step: 40, loss: 0.004623816814273596
step: 50, loss: 0.0012295832857489586
step: 60, loss: 0.0005591183435171843
step: 70, loss: 0.00030189973767846823
step: 80, loss: 0.0008787415572442114
step: 90, loss: 0.012797703966498375
step: 100, loss: 0.0010214063804596663
step: 110, loss: 0.0013491949066519737
step: 120, loss: 0.0010228789178654552
step: 130, loss: 0.0006880149012431502
step: 140, loss: 0.000570300268009305
step: 150, loss: 0.006826468743383884
step: 160, loss: 0.0023210851941257715
step: 170, loss: 0.003680839203298092
step: 180, loss: 0.0005463977577164769
step: 190, loss: 0.002008983166888356
step: 200, loss: 0.001470633433200419
step: 210, loss: 0.003773196367546916
/usr/local/lib/python3.7/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=0.5377358490566038, f1=0.46534653465346537,
best_f1=0.5406427221172023
```

```
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.")
2293it [00:06, 333.15it/s]
load_f1 = 0.5608465608465608
real_f1 = 0.5607142857142857
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, 256.77it/s]
```

2.1.5   Walmart-Amazon - Running the matcher

```python
# 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/
```

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: 0.8671017289161682
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.1665564924478531
step: 20, loss: 0.15412193536758423
step: 30, loss: 0.5110574960708618
step: 40, loss: 0.2525504529476166
step: 50, loss: 0.3123096227645874
step: 60, loss: 0.3612368404865265
step: 70, loss: 0.17517834901809692
step: 80, loss: 0.5278688073158264
step: 90, loss: 0.23655374348163605
step: 100, loss: 0.221385195851326
step: 110, loss: 0.23595726490020752
step: 120, loss: 0.4168330729007721
step: 130, loss: 0.3389187455177307
step: 140, loss: 0.32701781392097473
step: 150, loss: 0.2764228880405426
step: 160, loss: 0.21250471472740173
step: 170, loss: 0.38323622941970825
step: 180, loss: 0.3031391501426697
step: 190, loss: 0.12922878563404083
/usr/local/lib/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=0.4363636363636364, f1=0.45, best_f1=0.45
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2864669859409332
step: 10, loss: 0.03958810865879059
step: 20, loss: 0.07638322561979294
step: 30, loss: 0.14875490963459015
step: 40, loss: 0.4660681486129761
```

```
step: 50, loss: 0.3323560059070587
step: 60, loss: 0.14255017042160034
step: 70, loss: 0.20286774635314941
step: 80, loss: 0.1462486982345581
step: 90, loss: 0.11715566366910934
step: 100, loss: 0.3058686852455139
step: 110, loss: 0.12816981971263885
step: 120, loss: 0.27427515387535095
step: 130, loss: 0.15518446266651154
step: 140, loss: 0.2847101092338562
step: 150, loss: 0.024390241131186485
step: 160, loss: 0.026150230318307877
step: 170, loss: 0.28007471561431885
step: 180, loss: 0.1183619350194931
step: 190, loss: 0.18477903306484222
/usr/local/lib/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=0.7619047619047619, f1=0.7506849315068493,
best_f1=0.7506849315068493
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1451077163219452
step: 10, loss: 0.20539750158786774
step: 20, loss: 0.06385987251996994
step: 30, loss: 0.017605062574148178
step: 40, loss: 0.07564723491668701
step: 50, loss: 0.19745969772338867
step: 60, loss: 0.08665785193443298
step: 70, loss: 0.11807581782341003
step: 80, loss: 0.2898920774459839
step: 90, loss: 0.05492442846298218
step: 100, loss: 0.11050748825073242
step: 110, loss: 0.1290002167224884
step: 120, loss: 0.0426030196249485
step: 130, loss: 0.04020795598626137
step: 140, loss: 0.14890412986278534
step: 150, loss: 0.16566479206085205
step: 160, loss: 0.12384815514087677
step: 170, loss: 0.0409911572933197
step: 180, loss: 0.058012135326862335
step: 190, loss: 0.12822194397449493
/usr/local/lib/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=0.7780979827089338, f1=0.735042735042735,
best_f1=0.735042735042735
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10500335693359375
step: 10, loss: 0.053489167243242264
step: 20, loss: 0.1082548126578331
step: 30, loss: 0.019159456714987755
step: 40, loss: 0.0026095740031450987
step: 50, loss: 0.010084710083901882
step: 60, loss: 0.057323701679706573
step: 70, loss: 0.0035615686792880297
step: 80, loss: 0.05773366242647171
step: 90, loss: 0.0834057405591011
step: 100, loss: 0.015572546981275082
step: 110, loss: 0.02614096738398075
step: 120, loss: 0.04590379446744919
step: 130, loss: 0.29491758346557617
step: 140, loss: 0.08664058148860931
step: 150, loss: 0.025696996599435806
step: 160, loss: 0.03286924958229065
step: 170, loss: 0.01531215850263834
step: 180, loss: 0.305477112531662
step: 190, loss: 0.03077664040029049
/usr/local/lib/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=0.7899159663865546, f1=0.7492957746478874,
best_f1=0.7492957746478874
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02303825505077839
step: 10, loss: 0.011817356571555138
step: 20, loss: 0.08754213154315948
step: 30, loss: 0.012049692682921886
step: 40, loss: 0.036711182445287704

```
step: 50, loss: 0.07452409714460373
step: 60, loss: 0.036635927855968475
step: 70, loss: 0.002478846814483404
step: 80, loss: 0.007847769185900688
step: 90, loss: 0.02751060202717781
step: 100, loss: 0.1553022414445877
step: 110, loss: 0.0025925615336745977
step: 120, loss: 0.0011810056616672337
step: 130, loss: 0.03410087901353836
step: 140, loss: 0.016575196757912636
step: 150, loss: 0.00676250783726573
step: 160, loss: 0.014999769628047943
step: 170, loss: 0.0239087101072073
step: 180, loss: 0.043681684881448746
step: 190, loss: 0.23251061141490936
/usr/local/lib/python3.7/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=0.767123287671233, f1=0.7431693989071039,
best_f1=0.7492957746478874
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03560202941298485
step: 10, loss: 0.0012267784914001822
step: 20, loss: 0.011985265649855137
step: 30, loss: 0.0013894438743591309
step: 40, loss: 0.0034866377245634794
step: 50, loss: 0.0045856633223593235
step: 60, loss: 0.0032431979198008776
step: 70, loss: 0.07902000844478607
step: 80, loss: 0.16887709498405457
step: 90, loss: 0.019504928961396217
step: 100, loss: 0.001509528374299407
step: 110, loss: 0.01604163646697998
step: 120, loss: 0.0170753840357O652
step: 130, loss: 0.0031801846344023943
step: 140, loss: 0.027581579983234406
step: 150, loss: 0.09209093451499939
step: 160, loss: 0.004818504676222801
step: 170, loss: 0.1927081197500229
step: 180, loss: 0.008365850895643234
step: 190, loss: 0.0032808666583150625
/usr/local/lib/python3.7/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=0.7731958762886598, f1=0.7272727272727273,
best_f1=0.7492957746478874
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010947147384285927
step: 10, loss: 0.049243032932281494
step: 20, loss: 0.09822100400924683
step: 30, loss: 0.004212366417050362
step: 40, loss: 0.002121431054547429
step: 50, loss: 0.005276782903820276
step: 60, loss: 0.009558876045048237
step: 70, loss: 0.0011399866780266166
step: 80, loss: 0.004757499322295189
step: 90, loss: 0.0040769074112176895
step: 100, loss: 0.0046683852560818195
step: 110, loss: 0.007439236156642437
step: 120, loss: 0.02758067473769188
step: 130, loss: 0.029458062723279
step: 140, loss: 0.004932126961648464
step: 150, loss: 0.01334322988986969
step: 160, loss: 0.0024126581847667694
step: 170, loss: 0.015898428857326508
step: 180, loss: 0.09882564842700958
step: 190, loss: 0.025191055610775948
/usr/local/lib/python3.7/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=0.7700831024930748, f1=0.7597765363128492,
best_f1=0.7492957746478874
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002005777321755886
step: 10, loss: 0.0012128829257562757
step: 20, loss: 0.009786617942154408
step: 30, loss: 0.000898851256351918
step: 40, loss: 0.005738576874136925
step: 50, loss: 0.0007461786735802889
step: 60, loss: 0.0006059342413209379
step: 70, loss: 0.0007334846304729581
step: 80, loss: 0.007345159538090229
step: 90, loss: 0.0008456647046841681
step: 100, loss: 0.005146736279129982
step: 110, loss: 0.018180716782808304
step: 120, loss: 0.01309338677674532
step: 130, loss: 0.003148408606648445
step: 140, loss: 0.0021724223624914885
step: 150, loss: 0.0008426400600001216
step: 160, loss: 0.0006474041729234159
```

```
step: 170, loss: 0.0007416266016662121
step: 180, loss: 0.004954872187227011
step: 190, loss: 0.027141151949763298
/usr/local/lib/python3.7/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=0.7582417582417582, f1=0.7446808510638299,
best_f1=0.7492957746478874
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0975823849439621
step: 10, loss: 0.0002323416993021965
step: 20, loss: 0.000836840074043721
step: 30, loss: 0.1578834354877472
step: 40, loss: 0.023674819618463516
step: 50, loss: 0.0003746885631699115
step: 60, loss: 0.013784127309918404
step: 70, loss: 0.001136637874878943
step: 80, loss: 0.00069610576611101222
step: 90, loss: 0.00213751127012074
step: 100, loss: 0.0018057755660265684
step: 110, loss: 0.04696378484368324
step: 120, loss: 0.0045077865943312645
step: 130, loss: 0.0025278418324887753
step: 140, loss: 0.00017501821275800467
step: 150, loss: 0.0017130079213529825
step: 160, loss: 0.0002408513391856104
step: 170, loss: 0.004277459811419249
step: 180, loss: 0.007495474070310593
step: 190, loss: 0.00048166399938054383
/usr/local/lib/python3.7/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=0.7704485488126649, f1=0.7277628032345013,
best_f1=0.7492957746478874
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00026249553775414824
step: 10, loss: 0.0018351186299696565
step: 20, loss: 0.00047158703091554344
step: 30, loss: 0.0003256043419241905
step: 40, loss: 0.0002664352359715849
step: 50, loss: 0.00012779937242157757
step: 60, loss: 0.006110010668635368
step: 70, loss: 0.0004073083691764623
step: 80, loss: 0.00022238855308387429
```

```
step: 90, loss: 0.0009798100218176842
step: 100, loss: 0.0017861929954960942
step: 110, loss: 0.0005283541977405548
step: 120, loss: 0.000424577941885218
step: 130, loss: 0.001266227918677032
step: 140, loss: 0.0004579080268740654
step: 150, loss: 0.00018493906827643514
step: 160, loss: 0.00022895223810337484
step: 170, loss: 0.0034813066013157368
step: 180, loss: 0.000492543913424015
step: 190, loss: 0.002545716241002083
/usr/local/lib/python3.7/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=0.7480916030534353, f1=0.7443037974683544,
best_f1=0.7492957746478874
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.115910105407238
step: 10, loss: 0.0019564356189221144
step: 20, loss: 0.0008447892032563686
step: 30, loss: 0.13696229457855225
step: 40, loss: 0.0011317847529426217
step: 50, loss: 0.001248506479896605
step: 60, loss: 0.0016704094596207142
step: 70, loss: 0.0015435926616191864
step: 80, loss: 0.0006016377592459321
step: 90, loss: 0.0006300288951024413
step: 100, loss: 0.0006903348839841783
step: 110, loss: 0.0011067201849073172
step: 120, loss: 0.0003535756841301918
step: 130, loss: 0.001793782925233245
step: 140, loss: 0.0004271703655831516
step: 150, loss: 0.00029048597207292914
step: 160, loss: 0.00016135942132677883
step: 170, loss: 0.0005967036704532802
step: 180, loss: 0.002152959583327174
step: 190, loss: 0.00041211285861209035
/usr/local/lib/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=0.7922437673130193, f1=0.7444444444444444,
best_f1=0.7444444444444444
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00023311770928557962
step: 10, loss: 0.000851781980600208
step: 20, loss: 0.0004431398119777441
step: 30, loss: 0.0003653481835499406
step: 40, loss: 0.002592579461634159
step: 50, loss: 0.00033581958268769085
step: 60, loss: 0.0002739785995800048
step: 70, loss: 0.0005566274048760533
step: 80, loss: 0.00028804608155041933
step: 90, loss: 0.00047170071047730744
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 100, loss: 0.0005077846581116319
step: 110, loss: 0.00017947371816262603
step: 120, loss: 0.00028194847982376814
step: 130, loss: 0.20195147395133972
step: 140, loss: 0.002402404323220253
step: 150, loss: 0.00023379070626106113
step: 160, loss: 0.0016711963107809424
step: 170, loss: 0.00039249513065442443
step: 180, loss: 0.0002458913659211248
step: 190, loss: 0.0012413888471201062
/usr/local/lib/python3.7/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=0.7777777777777779, f1=0.7473684210526317,
best_f1=0.7444444444444444
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004058764025103301
step: 10, loss: 0.00038710585795534292
step: 20, loss: 0.00025545185781084
step: 30, loss: 0.00016335041436832398
step: 40, loss: 0.00011340115452185273
step: 50, loss: 0.00022426740906666964
step: 60, loss: 0.00040720796096138656
step: 70, loss: 0.00024818131350912154
step: 80, loss: 0.00016522414807695895
step: 90, loss: 0.00043294072384014726
step: 100, loss: 0.0004588933661580086
step: 110, loss: 0.00018245696264784783
step: 120, loss: 0.001209524110890925
step: 130, loss: 0.0004949024296365678
step: 140, loss: 0.00025617610663175583
step: 150, loss: 0.0004875287995673716
```

```
step: 160, loss: 0.00015650372370146215
step: 170, loss: 0.0002290929842274636
step: 180, loss: 0.0003713926416821778
step: 190, loss: 0.0015056314878165722
/usr/local/lib/python3.7/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=0.7903225806451613, f1=0.7405405405405405,
best_f1=0.7444444444444444
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00028518581530079246
step: 10, loss: 0.00017847676645033062
step: 20, loss: 0.0007052168948575854
step: 30, loss: 0.0005219726590439677
step: 40, loss: 0.000851778662763536
step: 50, loss: 0.0003646490804385394
step: 60, loss: 0.021570023149251938
step: 70, loss: 0.00023450521985068917
step: 80, loss: 6.589727126993239e-05
step: 90, loss: 0.00036443164572119713
step: 100, loss: 0.00019178660295438021
step: 110, loss: 0.00011468308366602287
step: 120, loss: 0.0002767247206065804
step: 130, loss: 0.0005068184109404683
step: 140, loss: 0.0001099331711884588
step: 150, loss: 0.0002067958121187985
step: 160, loss: 0.0001016178575810045
step: 170, loss: 0.00020613659580703825
step: 180, loss: 0.00024640001356601715
step: 190, loss: 0.0032721569295972586
/usr/local/lib/python3.7/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=0.7795698924731183, f1=0.7411444141689373,
best_f1=0.7444444444444444
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002706032828427851
step: 10, loss: 0.00011392379383323714
step: 20, loss: 0.00011943950084969401
step: 30, loss: 0.00014500202087219805
step: 40, loss: 0.0009274039184674621
step: 50, loss: 0.00021761705284006894
step: 60, loss: 9.652403969084844e-05
step: 70, loss: 0.0016895567532628775
```

```
step: 80, loss: 0.0001386008079862222
step: 90, loss: 0.13924188911914825
step: 100, loss: 0.0002516841923352331
step: 110, loss: 0.001014314591884613
step: 120, loss: 0.002054646611213684
step: 130, loss: 0.00026119413087144494
step: 140, loss: 0.000907577108591795
step: 150, loss: 0.0002941161801572889
step: 160, loss: 0.00018170363910030574
step: 170, loss: 0.00011520587577251717
step: 180, loss: 0.0002461190742906183
step: 190, loss: 0.0005695514264516532
/usr/local/lib/python3.7/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=0.7733333333333333, f1=0.7445652173913043,
best_f1=0.7444444444444444
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.")
2049it [00:09, 225.30it/s]
```

```
load_f1 = 0.7867036011080332
real_f1 = 0.7878787878787877
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, 256.75it/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 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: 0.8540818095207214
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.22437278926372528
step: 20, loss: 0.15094774961471558
step: 30, loss: 0.2363802194595337
step: 40, loss: 0.3111476004123688
step: 50, loss: 0.37436071038246155
step: 60, loss: 0.432976633310318
step: 70, loss: 0.30560633540153503
step: 80, loss: 0.26235491037368774
step: 90, loss: 0.40429309010505676
step: 100, loss: 0.23341277241706848
step: 110, loss: 0.18448051810264587
step: 120, loss: 0.5396993160247803
step: 130, loss: 0.416827529668808
step: 140, loss: 0.474312424659729
step: 150, loss: 0.11923571676015854
step: 160, loss: 0.3365550637245178
step: 170, loss: 0.19136768579483032
```

```
/usr/local/lib/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=0.5933014354066986, f1=0.5211267605633804,
best_f1=0.5211267605633804
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.41347789764404297
step: 10, loss: 0.08058462291955948
step: 20, loss: 0.3562828302383423
step: 30, loss: 0.1098581999540329
step: 40, loss: 0.1814471036195755
step: 50, loss: 0.2205204963684082
step: 60, loss: 0.17667317390441895
step: 70, loss: 0.2704121470451355
step: 80, loss: 0.0571659691631794
step: 90, loss: 0.15017183125019073
step: 100, loss: 0.1698225736618042
step: 110, loss: 0.23510116338729858
step: 120, loss: 0.1095418781042099
step: 130, loss: 0.07858660817146301
step: 140, loss: 0.12988318502902985
step: 150, loss: 0.16483847796916962
step: 160, loss: 0.2064262330532074
step: 170, loss: 0.16164231300354004
/usr/local/lib/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=0.727710843373494, f1=0.7330316742081447,
best_f1=0.7330316742081447
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05675715580582619
step: 10, loss: 0.03614141419529915
step: 20, loss: 0.04244533181190491
step: 30, loss: 0.2558319866657257
step: 40, loss: 0.023401658982038498
step: 50, loss: 0.13432161509990692
```

```
step: 60, loss: 0.25965309143066406
step: 70, loss: 0.11018113046884537
step: 80, loss: 0.2644636332988739
step: 90, loss: 0.08042339980602264
step: 100, loss: 0.02812981605529785
step: 110, loss: 0.14254900813102722
step: 120, loss: 0.007075728382915258
step: 130, loss: 0.17853303253650665
step: 140, loss: 0.014846570789813995
step: 150, loss: 0.05760766193270683
step: 160, loss: 0.03652231767773628
step: 170, loss: 0.11120161414146423
/usr/local/lib/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=0.7599067599067598, f1=0.7571115973741793,
best_f1=0.7571115973741793
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05835048481822014
step: 10, loss: 0.2230377346277237
step: 20, loss: 0.03717781603336334
step: 30, loss: 0.1708291620016098
step: 40, loss: 0.020011425018310547
step: 50, loss: 0.013608294539153576
step: 60, loss: 0.013565932400524616
step: 70, loss: 0.008681442588567734
step: 80, loss: 0.019221119582653046
step: 90, loss: 0.02307489700615406
step: 100, loss: 0.007131519261747599
step: 110, loss: 0.024402666836977005
step: 120, loss: 0.1727493852376938
step: 130, loss: 0.03317498788237572
step: 140, loss: 0.005434843711555004
step: 150, loss: 0.002768767299130559
step: 160, loss: 0.07139117270708084
step: 170, loss: 0.058759983628988266
/usr/local/lib/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=0.7673860911270983, f1=0.7670588235294118,
best_f1=0.7670588235294118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024553749710321426
step: 10, loss: 0.08499282598495483
step: 20, loss: 0.01416265219449997
step: 30, loss: 0.011542641557753086
step: 40, loss: 0.0035717191640287638
step: 50, loss: 0.013256956823170185
step: 60, loss: 0.04211015626788139
step: 70, loss: 0.007588736712932587
step: 80, loss: 0.014349974691867828
step: 90, loss: 0.029082829132676125
step: 100, loss: 0.005571593064814806
step: 110, loss: 0.12704254686832428
step: 120, loss: 0.011520477011799812
step: 130, loss: 0.0027315188199281693
step: 140, loss: 0.029609326273202896
step: 150, loss: 0.017542369663715363
step: 160, loss: 0.010445548221468925
step: 170, loss: 0.020467311143875122
/usr/local/lib/python3.7/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=0.7568922305764412, f1=0.7459207459207459,
best_f1=0.7670588235294118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013124573975801468
step: 10, loss: 0.020013723522424698
step: 20, loss: 0.0016640747198835015
step: 30, loss: 0.00295030465349555
step: 40, loss: 0.003908089362084866
step: 50, loss: 0.048122864216566086
step: 60, loss: 0.20751094818115234
step: 70, loss: 0.009404483251273632
step: 80, loss: 0.06445195525884628
step: 90, loss: 0.040905315428972244
step: 100, loss: 0.006062771659344435
step: 110, loss: 0.006758650299161673
step: 120, loss: 0.045333825051784515
step: 130, loss: 0.11250562220811844
step: 140, loss: 0.0019458773313090205
step: 150, loss: 0.14465442299842834
```

```
step: 160, loss: 0.1283802092075348
step: 170, loss: 0.0038952906616032124
/usr/local/lib/python3.7/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=0.742014742014742, f1=0.7868852459016392,
best_f1=0.7670588235294118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20931100845336914
step: 10, loss: 0.0017480513779446483
step: 20, loss: 0.02132115140557289
step: 30, loss: 0.03729984164237976
step: 40, loss: 0.0039442433044314384
step: 50, loss: 0.007862544618546963
step: 60, loss: 0.3363317847251892
step: 70, loss: 0.01885426789522171
step: 80, loss: 0.0044215042144060135
step: 90, loss: 0.016633978113532066
step: 100, loss: 0.007754234131425619
step: 110, loss: 0.0004989327280782163
step: 120, loss: 0.18807139992713928
step: 130, loss: 0.0724082887172699
step: 140, loss: 0.006276663392782211
step: 150, loss: 0.02173951454460621
step: 160, loss: 0.0027968953363597393
step: 170, loss: 0.2250118851661682
/usr/local/lib/python3.7/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=0.7450110864745012, f1=0.7291242362525456,
best_f1=0.7670588235294118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018640143796801567
step: 10, loss: 0.003482986707240343
step: 20, loss: 0.00853261724114418
step: 30, loss: 0.005941241513937712
step: 40, loss: 0.0006533594569191337
step: 50, loss: 0.000700717733707279
step: 60, loss: 0.02617490105330944
step: 70, loss: 0.19399720430374146
step: 80, loss: 0.001356596709229052
step: 90, loss: 0.014251376502215862
step: 100, loss: 0.019038071855902672
step: 110, loss: 0.05345357954502106
```

```
step: 120, loss: 0.048229772597551346
step: 130, loss: 0.00765112042427063
step: 140, loss: 0.009667919017374516
step: 150, loss: 0.03728647902607918
step: 160, loss: 0.01609189622104168
step: 170, loss: 0.013033575378358364
/usr/local/lib/python3.7/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=0.756152125279642, f1=0.7386609071274297,
best_f1=0.7670588235294118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0024065582547336817
step: 10, loss: 0.019327370449900627
step: 20, loss: 0.0004982841201126575
step: 30, loss: 0.026782667264342308
step: 40, loss: 0.013961476273834705
step: 50, loss: 0.0055048903450369835
step: 60, loss: 0.00043902284232899547
step: 70, loss: 0.014967831782996655
step: 80, loss: 0.031094757840037346
step: 90, loss: 0.003318179165944457
step: 100, loss: 0.022568345069885254
step: 110, loss: 0.0007507543778046966
step: 120, loss: 0.013419970870018005
step: 130, loss: 0.00247801817022264
step: 140, loss: 0.000890677678398788
step: 150, loss: 0.029095152392983437
step: 160, loss: 0.014947650954127312
step: 170, loss: 0.014255558140575886
/usr/local/lib/python3.7/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=0.7386091127098322, f1=0.7424593967517402,
best_f1=0.7670588235294118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00968723464757204
step: 10, loss: 0.00418986938893795
step: 20, loss: 0.0005569179193116724
step: 30, loss: 0.0009625514503568411
step: 40, loss: 0.0017010787269100547
step: 50, loss: 0.0010991047602146864
step: 60, loss: 0.000666609441395849
step: 70, loss: 0.000362083112122491
```

```
step: 80, loss: 0.0009847708279266953
step: 90, loss: 0.09830542653799057
step: 100, loss: 0.003308109473437071
step: 110, loss: 0.0010544054675847292
step: 120, loss: 0.015463822521269321
step: 130, loss: 0.005734439007937908
step: 140, loss: 0.005096630193293095
step: 150, loss: 0.0007202327251434326
step: 160, loss: 0.006070148665457964
step: 170, loss: 0.00032609421759843826
/usr/local/lib/python3.7/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=0.7506426735218509, f1=0.7696078431372548,
best_f1=0.7670588235294118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006460136268287897
step: 10, loss: 0.10899823904037476
step: 20, loss: 0.05519286170601845
step: 30, loss: 0.006682416889816523
step: 40, loss: 0.056995321065187454
step: 50, loss: 0.02039647474884987
step: 60, loss: 0.0020553343929350376
step: 70, loss: 0.002313896780833602
step: 80, loss: 0.017809828743338585
step: 90, loss: 0.0030979386065155268
step: 100, loss: 0.0010669566690921783
step: 110, loss: 0.003619071561843157
step: 120, loss: 0.0013539927313104272
step: 130, loss: 0.0007469283882528543
step: 140, loss: 0.005458103958517313
step: 150, loss: 0.00047324964543804526
step: 160, loss: 0.0001286475162487477
step: 170, loss: 0.06736584007740021
/usr/local/lib/python3.7/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=0.7227722772277227, f1=0.7592592592592593,
best_f1=0.7670588235294118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004160830285400152
step: 10, loss: 0.0037971187848597765
step: 20, loss: 0.0274571031332016
step: 30, loss: 0.0032917119096964598
```

```
step: 40, loss: 0.0001476956531405449
step: 50, loss: 0.0010319288121536374
step: 60, loss: 0.00023877430066931993
step: 70, loss: 0.030874967575073242
step: 80, loss: 0.04039711132645607
step: 90, loss: 0.0006713818875141442
step: 100, loss: 0.0001899121270980686
step: 110, loss: 0.0007467342657037079
step: 120, loss: 0.00103804434183985
step: 130, loss: 0.0005013610934838653
step: 140, loss: 0.00022386440832633525
step: 150, loss: 0.0013094430323690176
step: 160, loss: 0.006010373588651419
step: 170, loss: 0.007153376005589962
/usr/local/lib/python3.7/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=0.7391304347826086, f1=0.7500000000000001,
best_f1=0.7670588235294118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019374852126922991
step: 10, loss: 0.03955576941370964
step: 20, loss: 0.01432271022349596
step: 30, loss: 0.00043242188985459507
step: 40, loss: 0.0002450906322337687
step: 50, loss: 0.004843046423047781
step: 60, loss: 0.00048411317402496934
step: 70, loss: 0.005756321828812361
step: 80, loss: 0.001356065971776843
step: 90, loss: 0.0003312610206194222
step: 100, loss: 0.00018806253501679748
step: 110, loss: 0.00026448097196407616
step: 120, loss: 0.0006075128330849111
step: 130, loss: 0.00019641040125861764
step: 140, loss: 0.00015544208872597665
step: 150, loss: 0.12629659473896027
step: 160, loss: 0.005503907799720764
step: 170, loss: 0.0025662914849817753
/usr/local/lib/python3.7/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=0.745679012345679, f1=0.7540229885057472,
best_f1=0.7670588235294118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.01123399380594492
step: 10, loss: 0.0003636575420387089
step: 20, loss: 0.008402814157307148
step: 30, loss: 0.013243262656033039
step: 40, loss: 0.0006327711162157357
step: 50, loss: 0.0036128093488514423
step: 60, loss: 0.0023565383162349463
step: 70, loss: 0.001153046265244484
step: 80, loss: 0.000545334187336266
step: 90, loss: 0.0003369661862961948
step: 100, loss: 0.00017950829351320863
step: 110, loss: 0.000130328320665285
step: 120, loss: 0.005336184054613113
step: 130, loss: 0.0012132578995078802
step: 140, loss: 0.0002376218471908942
step: 150, loss: 0.0008994521922431886
step: 160, loss: 0.044942084699869156
step: 170, loss: 0.0002633194380905479
/usr/local/lib/python3.7/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=0.7439613526570048, f1=0.749425287356322,
best_f1=0.7670588235294118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005590965738520026
step: 10, loss: 0.00981514435261488
step: 20, loss: 0.04344929754734039
step: 30, loss: 0.020113512873649597
step: 40, loss: 0.0002153862442355603
step: 50, loss: 0.0005496490048244596
step: 60, loss: 0.00013668983592651784
step: 70, loss: 0.0023047826252877712
step: 80, loss: 0.0002583638997748494
step: 90, loss: 0.00010454311995999888
step: 100, loss: 0.00026083318516612053
step: 110, loss: 0.00018926538177765906
step: 120, loss: 0.0009658602648414671
step: 130, loss: 0.004410077817738056
step: 140, loss: 0.0003285915299784392
step: 150, loss: 0.010251648724079132
step: 160, loss: 0.00012460630387067795
step: 170, loss: 0.00023534748470410705
/usr/local/lib/python3.7/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=0.7427184466019418, f1=0.7546296296296297,
```

```
best_f1=0.7670588235294118
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.")
1916it [00:04, 416.88it/s]
load_f1 = 0.7668161434977578
real_f1 = 0.7627494456762749
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 [00:12, 351.16it/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 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: 0.8031815886497498
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4813767671585083
step: 20, loss: 0.6112840175628662
step: 30, loss: 0.49515005946159363
step: 40, loss: 0.3554045557975769
step: 50, loss: 0.23356151580810547
step: 60, loss: 0.14924457669258118
step: 70, loss: 0.28943932056427
step: 80, loss: 0.2237376868724823
step: 90, loss: 0.16417139768600464
step: 100, loss: 0.17678911983966827
step: 110, loss: 0.17696361243724823
step: 120, loss: 0.022434769198298454
step: 130, loss: 0.01561364158987999
step: 140, loss: 0.1721566915512085
step: 150, loss: 0.10282072424888611
step: 160, loss: 0.23137351870536804
step: 170, loss: 0.011920680291950703
step: 180, loss: 0.0115338871255517
step: 190, loss: 0.10755425691604614
step: 200, loss: 0.009535178542137146
step: 210, loss: 0.015113834291696548
step: 220, loss: 0.0367441289126873
step: 230, loss: 0.010079648345708847
/usr/local/lib/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=0.9555555555555556, f1=0.9661399548532732,
best_f1=0.9661399548532732
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05351642519235611
step: 10, loss: 0.040370963513851166
step: 20, loss: 0.005532668437808752
step: 30, loss: 0.010890709236264229
step: 40, loss: 0.047078834869623184
step: 50, loss: 0.017770634964108467
step: 60, loss: 0.08997061103582382
step: 70, loss: 0.06837119907140732
step: 80, loss: 0.008907823823392391
step: 90, loss: 0.006863829679787159
step: 100, loss: 0.11115202307701111
step: 110, loss: 0.05168914422392845
step: 120, loss: 0.004746549762785435
step: 130, loss: 0.015041636303067207
step: 140, loss: 0.1905592381954193
step: 150, loss: 0.03150033578276634
step: 160, loss: 0.022883731871843338
step: 170, loss: 0.0802680253982544
step: 180, loss: 0.0436708889901638
step: 190, loss: 0.12132642418146133
step: 200, loss: 0.1309024542570114
step: 210, loss: 0.08597066253423691
step: 220, loss: 0.0025436978321522474
step: 230, loss: 0.017183156684041023
/usr/local/lib/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=0.9683972911963882, 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: 0.047748398035764694
step: 10, loss: 0.01797872968018055
step: 20, loss: 0.007273358292877674
step: 30, loss: 0.01733456738293171
step: 40, loss: 0.008001619949936867
step: 50, loss: 0.002444571815431118
step: 60, loss: 0.18559904396533966
step: 70, loss: 0.0213174968957901
step: 80, loss: 0.010325687006115913
```

```
step: 90, loss: 0.06489843130111694
step: 100, loss: 0.003394901752471924
step: 110, loss: 0.005392767488956451
step: 120, loss: 0.039755772799253464
step: 130, loss: 0.0031692013144493103
step: 140, loss: 0.001268765889108181
step: 150, loss: 0.015517947264015675
step: 160, loss: 0.0016382780158892274
step: 170, loss: 0.006738265044987202
step: 180, loss: 0.002595462603494525
step: 190, loss: 0.00591524550691247
step: 200, loss: 0.015553424134850502
step: 210, loss: 0.0012234478490427136
step: 220, loss: 0.006941284518688917
step: 230, loss: 0.1537501960992813
/usr/local/lib/python3.7/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=0.9557522123893805, f1=0.968609865470852,
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: 0.01630922220647335
step: 10, loss: 0.003692896571010351
step: 20, loss: 0.04850317910313606
step: 30, loss: 0.0053313123062253
step: 40, loss: 0.059073738753795624
step: 50, loss: 0.0016372054815292358
step: 60, loss: 0.00045959404087625444
step: 70, loss: 0.05340257287025452
step: 80, loss: 0.1889621764421463
step: 90, loss: 0.05113793909549713
step: 100, loss: 0.006433209870010614
step: 110, loss: 0.00625282758846879
step: 120, loss: 0.0030079081188887358
step: 130, loss: 0.011230912990868092
step: 140, loss: 0.006290575489401817
step: 150, loss: 0.004038363229483366
step: 160, loss: 0.002952399430796504
step: 170, loss: 0.0019252472557127476
step: 180, loss: 0.13691923022270203
step: 190, loss: 0.0032671133521944284
step: 200, loss: 0.0011006074491888285
step: 210, loss: 0.0011027175933122635
step: 220, loss: 0.008451727218925953
step: 230, loss: 0.00040610763244330883
/usr/local/lib/python3.7/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=0.9634551495016611, f1=0.9698324022346367,
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: 0.0006326555740088224
step: 10, loss: 0.0009163108770735562
step: 20, loss: 0.0004928719718009233
step: 30, loss: 0.0005131209036335349
step: 40, loss: 0.0005523822037503123
step: 50, loss: 0.00027665222296491265
step: 60, loss: 0.0029211246874183416
step: 70, loss: 0.0002809281868394464
step: 80, loss: 0.015252163633704185
step: 90, loss: 0.00046112522250041366
step: 100, loss: 0.0011295019648969173
step: 110, loss: 0.00042011452023871243
step: 120, loss: 0.0043820166029036045
step: 130, loss: 0.12997356057167053
step: 140, loss: 0.0006238357746042311
step: 150, loss: 0.002655575517565012
step: 160, loss: 0.004534402396529913
step: 170, loss: 0.10391827672719955
step: 180, loss: 0.0019452630076557398
step: 190, loss: 0.0003013864334207028
step: 200, loss: 0.0008776164031587541
step: 210, loss: 0.0004626682202797383
step: 220, loss: 0.0014201318845152855
step: 230, loss: 0.04906252771615982
/usr/local/lib/python3.7/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=0.9662162162162162, f1=0.9614512471655329,
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: 0.03003085032105446
step: 10, loss: 0.10444437712430954
step: 20, loss: 0.0017075539799407125
step: 30, loss: 0.014998525381088257
step: 40, loss: 0.009036194533109665
step: 50, loss: 0.052585795521736145
step: 60, loss: 0.0015757132787257433
step: 70, loss: 0.0006278425571508706
step: 80, loss: 0.0444958470761776
```

```
step: 90, loss: 0.00031644132104702294
step: 100, loss: 0.0004008904506918043
step: 110, loss: 0.009054920636117458
step: 120, loss: 0.001030179439112544
step: 130, loss: 0.035474736243486404
step: 140, loss: 0.02411539852619171
step: 150, loss: 0.0005442266701720655
step: 160, loss: 0.009973128326237202
step: 170, loss: 0.00024250622664112598
step: 180, loss: 0.00041360710747539997
step: 190, loss: 0.007112609222531319
step: 200, loss: 0.0024373044725507498
step: 210, loss: 0.0009870901703834534
step: 220, loss: 0.0002903858257923275
step: 230, loss: 0.0008625152986496687
/usr/local/lib/python3.7/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=0.9573770491803277, f1=0.9595628415300547,
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: 0.16016855835914612
step: 10, loss: 0.0004948978894390166
step: 20, loss: 0.03879480063915253
step: 30, loss: 0.00734697375446558
step: 40, loss: 0.008257164619863033
step: 50, loss: 0.001327390200458467
step: 60, loss: 0.01897593028843403
step: 70, loss: 0.0006538930465467274
step: 80, loss: 0.0001306606864091009
step: 90, loss: 0.0017833681777119637
step: 100, loss: 0.010356638580560684
step: 110, loss: 0.0005396487540565431
step: 120, loss: 0.001143446541391313
step: 130, loss: 0.0006935861892998219
step: 140, loss: 0.00019010924734175205
step: 150, loss: 0.0027186910156160593
step: 160, loss: 0.0014524203725159168
step: 170, loss: 0.0003209603310097009
step: 180, loss: 0.0003999084874521941
step: 190, loss: 0.09411289542913437
step: 200, loss: 0.0007766729104332626
step: 210, loss: 0.00031050515826791525
step: 220, loss: 0.06002353876829147
step: 230, loss: 0.017539335414767265
/usr/local/lib/python3.7/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=0.9641255605381166, f1=0.9707865168539327,
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: 0.0005131885991431773
step: 10, loss: 0.0017740780021995306
step: 20, loss: 0.010706093162298203
step: 30, loss: 0.00039023137651383877
step: 40, loss: 0.0007718075066804886
step: 50, loss: 0.0006072070100344718
step: 60, loss: 0.0006889545475132763
step: 70, loss: 0.0014090656768530607
step: 80, loss: 0.00035095837665721774
step: 90, loss: 0.0010714292293414474
step: 100, loss: 0.0015116077847778797
step: 110, loss: 0.00023967481683939695
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 120, loss: 0.0004949663998559117
step: 130, loss: 0.0007862279308028519
step: 140, loss: 9.993357525672764e-05
step: 150, loss: 0.00014138789265416563
step: 160, loss: 0.0016123858513310552
step: 170, loss: 0.00014230019587557763
step: 180, loss: 0.002204868709668517
step: 190, loss: 6.209465936990455e-05
step: 200, loss: 0.003075035521760583
step: 210, loss: 0.0005336907343007624
step: 220, loss: 0.00013413782289717346
step: 230, loss: 0.10432293266057968
/usr/local/lib/python3.7/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=0.9585666293393057, f1=0.9650507328072153,
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: 8.2414124335628e-05
step: 10, loss: 0.0035333866253495216
step: 20, loss: 0.0009366390295532671
step: 30, loss: 0.0002468073507770896
step: 40, loss: 0.03867335990071297
step: 50, loss: 0.0014443048276007175
step: 60, loss: 0.0029202657751739025
step: 70, loss: 0.0003701056120917201
```

```
step: 80, loss: 0.007579816970974207
step: 90, loss: 0.0022925077937543392
step: 100, loss: 0.002134721027687192
step: 110, loss: 7.44485150789842e-05
step: 120, loss: 0.025225579738616943
step: 130, loss: 0.0001855466398410499
step: 140, loss: 0.045705363154411316
step: 150, loss: 9.748958109412342e-05
step: 160, loss: 0.13493762910366058
step: 170, loss: 0.00011822429223684594
step: 180, loss: 0.0016450616531074047
step: 190, loss: 0.00019130902364850044
step: 200, loss: 0.00041461718501523137
step: 210, loss: 0.00043890863889828324
step: 220, loss: 0.0005696816951967776
step: 230, loss: 0.010649314150214195
/usr/local/lib/python3.7/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=0.9652855543113102, f1=0.9707865168539327,
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: 0.0019878826569765806
step: 10, loss: 0.0007386479410342872
step: 20, loss: 0.0018312516622245312
step: 30, loss: 0.0004310766817070544
step: 40, loss: 0.0003058917645830661
step: 50, loss: 0.0036217730958014727
step: 60, loss: 0.009840883314609528
step: 70, loss: 0.0019370605004951358
step: 80, loss: 0.00017740776820573956
step: 90, loss: 0.00016833505651447922
step: 100, loss: 0.00037246564170345664
step: 110, loss: 0.0007021305500529706
step: 120, loss: 0.016423363238573074
step: 130, loss: 0.0001869949046522379
step: 140, loss: 0.004888170398771763
step: 150, loss: 0.00010214576468570158
step: 160, loss: 0.00011719930625986308
step: 170, loss: 0.00012989465903956443
step: 180, loss: 0.0001071263977792114
step: 190, loss: 0.0003995893057435751
step: 200, loss: 7.039100455585867e-05
step: 210, loss: 8.724658255232498e-05
step: 220, loss: 0.0028434544801712036
step: 230, loss: 0.12896206974983215
```

```
/usr/local/lib/python3.7/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=0.9633740288568259, f1=0.9654403567447045,
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: 0.0004100594378542155
step: 10, loss: 0.00047742482274770737
step: 20, loss: 6.378758553182706e-05
step: 30, loss: 0.00020593992667272687
step: 40, loss: 0.04751238226890564
step: 50, loss: 0.0012851201463490725
step: 60, loss: 0.01707102730870247
step: 70, loss: 0.0003635923203546554
step: 80, loss: 0.0004429617547430098
step: 90, loss: 0.014520091004669666
step: 100, loss: 0.009231911040842533
step: 110, loss: 0.0011064520804211497
step: 120, loss: 0.0004060290229972452
step: 130, loss: 5.133444210514426e-05
step: 140, loss: 0.00012410996671460688
step: 150, loss: 6.130022666184232e-05
step: 160, loss: 0.00011814099707407877
step: 170, loss: 0.001412575482390821
step: 180, loss: 0.0023790837731212378
step: 190, loss: 0.00018544006161391735
step: 200, loss: 8.768635598244146e-05
step: 210, loss: 0.00010912663856288418
step: 220, loss: 0.0008801035583019257
step: 230, loss: 0.0013171923346817493
/usr/local/lib/python3.7/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=0.9652855543113102, f1=0.9685393258426966,
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: 0.010829536244273186
step: 10, loss: 0.00020420164219103754
step: 20, loss: 5.901912663830444e-05
step: 30, loss: 0.06325628608465195
step: 40, loss: 0.0002322602813364938
step: 50, loss: 0.00017809250857681036
step: 60, loss: 5.289609180181287e-05
step: 70, loss: 0.029641173779964447
```

```
step: 80, loss: 9.132386185228825e-05
step: 90, loss: 0.00016158205107785761
step: 100, loss: 3.9575159462401643e-05
step: 110, loss: 0.0029962449334561825
step: 120, loss: 0.00011290621478110552
step: 130, loss: 0.00018219499906990677
step: 140, loss: 8.799156785244122e-05
step: 150, loss: 0.0003000142169184983
step: 160, loss: 0.002957182703539729
step: 170, loss: 0.00010786568600451574
step: 180, loss: 5.5709788284730166e-05
step: 190, loss: 5.0014237785944715e-05
step: 200, loss: 8.957940008258447e-05
step: 210, loss: 0.00020801181381102651
step: 220, loss: 0.0006241020164452493
step: 230, loss: 0.00010580387606751174
/usr/local/lib/python3.7/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=0.9655937846836848, f1=0.967670011148272,
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: 6.362302519846708e-05
step: 10, loss: 4.770253872266039e-05
step: 20, loss: 7.712707156315446e-05
step: 30, loss: 0.0006950847455300391
step: 40, loss: 0.0020728912204504013
step: 50, loss: 0.00019515662279445678
step: 60, loss: 0.00015435802924912423
step: 70, loss: 7.548568284497082e-05
step: 80, loss: 2.8318314434727654e-05
step: 90, loss: 3.097080116276629e-05
step: 100, loss: 6.277690408751369e-05
step: 110, loss: 6.965642387513071e-05
step: 120, loss: 4.3732488848036155e-05
step: 130, loss: 0.0002296615275554359
step: 140, loss: 0.03664559870958328
step: 150, loss: 0.001588733633980155
step: 160, loss: 5.439157030195929e-05
step: 170, loss: 5.3618419769918546e-05
step: 180, loss: 0.00023358347243629396
step: 190, loss: 0.004686116240918636
step: 200, loss: 0.00014124074368737638
step: 210, loss: 0.00017170664796140045
step: 220, loss: 0.00015293697651941334
step: 230, loss: 3.549886241671629e-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=0.9644444444444443, f1=0.9665924276169264,
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: 4.19142052123788874e-05
step: 10, loss: 2.697380659810733e-05
step: 20, loss: 0.00029038041247986257
step: 30, loss: 6.578680768143386e-05
step: 40, loss: 6.740553362760693e-05
step: 50, loss: 7.999671652214602e-05
step: 60, loss: 4.7827088565099984e-05
step: 70, loss: 5.2514726121444255e-05
step: 80, loss: 8.417760545853525e-05
step: 90, loss: 0.00013146179844625294
step: 100, loss: 0.0010193852940574288
step: 110, loss: 0.00011229006486246362
step: 120, loss: 0.0001653884828556329
step: 130, loss: 0.00013669503096025437
step: 140, loss: 5.958710244158283e-05
step: 150, loss: 7.023661601124331e-05
step: 160, loss: 2.7461848731036298e-05
step: 170, loss: 9.080610470846295e-05
step: 180, loss: 0.00017154852685052902
step: 190, loss: 0.00015103619080036879
step: 200, loss: 3.4740736737148836e-05
step: 210, loss: 0.00016365337069146335
step: 220, loss: 0.0003704259288497269
step: 230, loss: 2.1773834305349737e-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=0.9643652561247216, f1=0.96875, 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: 5.792137744720094e-05
step: 10, loss: 0.00012438699195627123
step: 20, loss: 7.404419011436403e-05
step: 30, loss: 7.465860107913613e-05
step: 40, loss: 0.00044415571028366685
step: 50, loss: 0.0001608438033144921
step: 60, loss: 4.1790113755268976e-05
step: 70, loss: 7.512814772780985e-05
step: 80, loss: 0.00010296942491549999
```

```
step: 90, loss: 3.63523795385845e-05
step: 100, loss: 6.820596900070086e-05
step: 110, loss: 9.954375855159014e-05
step: 120, loss: 8.153558883350343e-05
step: 130, loss: 7.678406109334901e-05
step: 140, loss: 0.009887050837278366
step: 150, loss: 7.00534219504334e-05
step: 160, loss: 0.00013805553317070007
step: 170, loss: 4.542538590612821e-05
step: 180, loss: 0.00013249361654743552
step: 190, loss: 0.0009954373817890882
step: 200, loss: 6.3079729443416e-05
step: 210, loss: 0.0018811726476997137
step: 220, loss: 0.0004971736343577504
step: 230, loss: 9.233728633262217e-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=0.9645232815964524, f1=0.9655937846836848,
best_f1=0.9727891156462585
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 [00:08, 306.32it/s]
load_f1 = 0.9683257918552037
real_f1 = 0.967305524239008
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 [00:12, 340.93it/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 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: 0.7962470054626465
step: 10, loss: 0.41413426399230957
step: 20, loss: 0.4966309368610382
step: 30, loss: 0.4453980028629303
step: 40, loss: 0.4279820919036865
step: 50, loss: 0.16268005967140198
step: 60, loss: 0.32245033979415894
step: 70, loss: 0.18316099047660828
step: 80, loss: 0.11515167355537415
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 90, loss: 0.2892065644264221
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 100, loss: 0.1905546933412552
step: 110, loss: 0.10803408175706863
step: 120, loss: 0.06314537674188614
step: 130, loss: 0.02961243875324726
step: 140, loss: 0.36220937967300415
step: 150, loss: 0.057352542877197266
```

```
step: 160, loss: 0.1381523609161377
step: 170, loss: 0.24491511285305023
step: 180, loss: 0.14847375452518463
step: 190, loss: 0.0663207471370697
step: 200, loss: 0.09945371747016907
step: 210, loss: 0.13690932095050812
step: 220, loss: 0.20288731157779694
step: 230, loss: 0.08891886472702026
step: 240, loss: 0.09239636361598969
step: 250, loss: 0.06093674525618553
step: 260, loss: 0.008511165156960487
step: 270, loss: 0.015208573080599308
step: 280, loss: 0.14306463301181793
step: 290, loss: 0.05745520442724228
step: 300, loss: 0.15479376912117004
step: 310, loss: 0.09848842769861221
step: 320, loss: 0.039314549416303635
step: 330, loss: 0.05393955484032631
step: 340, loss: 0.15025800466537476
step: 350, loss: 0.10320988297462463
step: 360, loss: 0.05614447966217995
step: 370, loss: 0.10315846651792526
step: 380, loss: 0.14796945452690125
step: 390, loss: 0.02278665080666542
step: 400, loss: 0.0179084874689579
step: 410, loss: 0.039807360619306564
step: 420, loss: 0.02078075334429741
step: 430, loss: 0.0781085416674614
step: 440, loss: 0.19999979436397552
step: 450, loss: 0.08932803571224213
step: 460, loss: 0.022887974977493286
step: 470, loss: 0.35481590032577515
step: 480, loss: 0.20150622725486755
step: 490, loss: 0.06096741557121277
step: 500, loss: 0.006134131457656622
step: 510, loss: 0.1095295175909996
step: 520, loss: 0.07812528312206268
step: 530, loss: 0.0791754424571991
/usr/local/lib/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=0.920796665122742, f1=0.912540490513651,
best_f1=0.912540490513651
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09983545541763306
step: 10, loss: 0.13940848410129547
step: 20, loss: 0.12338301539421082
step: 30, loss: 0.19506269693374634
step: 40, loss: 0.015577663667500019
step: 50, loss: 0.038174550980329514
step: 60, loss: 0.19160811603069305
step: 70, loss: 0.06908886134624481
step: 80, loss: 0.028035571798682213
step: 90, loss: 0.031634192913770676
step: 100, loss: 0.19801096618175507
step: 110, loss: 0.03452582284808159
step: 120, loss: 0.1430262178182602
step: 130, loss: 0.03835175186395645
step: 140, loss: 0.03846530616283417
step: 150, loss: 0.07423973083496094
step: 160, loss: 0.05916866287589073
step: 170, loss: 0.12032756954431534
step: 180, loss: 0.007443315349519253
step: 190, loss: 0.029781214892864227
step: 200, loss: 0.0806079581379904
step: 210, loss: 0.10866621136665344
step: 220, loss: 0.10596293956041336
step: 230, loss: 0.017163433134555817
step: 240, loss: 0.1564338505268097
step: 250, loss: 0.15474280714988708
step: 260, loss: 0.010148736648261547
step: 270, loss: 0.04166710376739502
step: 280, loss: 0.20514114201068878
step: 290, loss: 0.01847982592880726
step: 300, loss: 0.05613669380545616
step: 310, loss: 0.04643973708152771
step: 320, loss: 0.18941733241081238
step: 330, loss: 0.10057423263788223
step: 340, loss: 0.012728424742817879
step: 350, loss: 0.08001337945461273
step: 360, loss: 0.024324199184775352
step: 370, loss: 0.017124108970165253
step: 380, loss: 0.06841826438903809
step: 390, loss: 0.035349857062101364
step: 400, loss: 0.1282777041196823
step: 410, loss: 0.0016374506521970034
step: 420, loss: 0.06244245171546936
step: 430, loss: 0.016188394278287888
step: 440, loss: 0.011879000812768936
step: 450, loss: 0.02441897802054882
```

```
step: 460, loss: 0.24229294061660767
step: 470, loss: 0.008779564872384071
step: 480, loss: 0.1879010647535324
step: 490, loss: 0.17763082683086395
step: 500, loss: 0.010461869649589062
step: 510, loss: 0.033420100808143616
step: 520, loss: 0.06293265521526337
step: 530, loss: 0.22924229502677917
/usr/local/lib/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=0.9282428702851886, f1=0.9262295081967213,
best_f1=0.9262295081967213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013023046776652336
step: 10, loss: 0.018773313611745834
step: 20, loss: 0.25690361857414246
step: 30, loss: 0.23403799533843994
step: 40, loss: 0.007960306480526924
step: 50, loss: 0.014087947085499763
step: 60, loss: 0.0246998630464077
step: 70, loss: 0.004840884357690811
step: 80, loss: 0.006040440406650305
step: 90, loss: 0.06931172311306
step: 100, loss: 0.010816158726811409
step: 110, loss: 0.0192562248557806
step: 120, loss: 0.06453735381364822
step: 130, loss: 0.0024241330102086067
step: 140, loss: 0.08210291713476181
step: 150, loss: 0.0191316157579422
step: 160, loss: 0.005327098537236452
step: 170, loss: 0.01760408841073513
step: 180, loss: 0.011878897435963154
step: 190, loss: 0.015736952424049377
step: 200, loss: 0.11207546293735504
step: 210, loss: 0.1350180208683014
step: 220, loss: 0.01708812639117241
step: 230, loss: 0.020388299599289894
step: 240, loss: 0.047366008162498474
step: 250, loss: 0.027604490518569946
step: 260, loss: 0.004893232136964798
step: 270, loss: 0.0033907569013535976
```

```
step: 280, loss: 0.17283643782138824
step: 290, loss: 0.03688560426235199
step: 300, loss: 0.0855773389339447
step: 310, loss: 0.19599775969982147
step: 320, loss: 0.17767302691936493
step: 330, loss: 0.010651066899299622
step: 340, loss: 0.0022849191445857286
step: 350, loss: 0.027914343401789665
step: 360, loss: 0.04045380651950836
step: 370, loss: 0.017313024029135704
step: 380, loss: 0.03922868147492409
step: 390, loss: 0.0040922630578279495
step: 400, loss: 0.015561224892735481
step: 410, loss: 0.015545411966741085
step: 420, loss: 0.04549535736441612
step: 430, loss: 0.004659298807382584
step: 440, loss: 0.005717126186937094
step: 450, loss: 0.0799834281206131
step: 460, loss: 0.11534512042999268
step: 470, loss: 0.0056706867180764675
step: 480, loss: 0.02359478361904621
step: 490, loss: 0.0070288535207509995
step: 500, loss: 0.12177810072898865
step: 510, loss: 0.05772976949810982
step: 520, loss: 0.0336775928735733
step: 530, loss: 0.03765485808253288
/usr/local/lib/python3.7/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=0.9237209302325581, f1=0.921999065857076,
best_f1=0.9262295081967213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0024604264181107283
step: 10, loss: 0.005648910999298096
step: 20, loss: 0.0021901049185544252
step: 30, loss: 0.08245068788528442
step: 40, loss: 0.00737807946279645
step: 50, loss: 0.026872918009757996
step: 60, loss: 0.0012403426226228476
step: 70, loss: 0.0060899690724909306
step: 80, loss: 0.05676192045211792
step: 90, loss: 0.021264802664518356
step: 100, loss: 0.23849953711032867
step: 110, loss: 0.025147004052996635
step: 120, loss: 0.0010383299086242914
step: 130, loss: 0.00916843581944704
```

```
step: 140, loss: 0.003559097880497575
step: 150, loss: 0.0018337341025471687
step: 160, loss: 0.017573274672031403
step: 170, loss: 0.04169126972556114
step: 180, loss: 0.006897349376231432
step: 190, loss: 0.00551247363910079
step: 200, loss: 0.0006287034484557807
step: 210, loss: 0.010732179507613182
step: 220, loss: 0.04299949109554291
step: 230, loss: 0.1402028650045395
step: 240, loss: 0.0018585423240438104
step: 250, loss: 0.0022658519446849823
step: 260, loss: 0.1530042588710785
step: 270, loss: 0.030367424711585045
step: 280, loss: 0.004556439816951752
step: 290, loss: 0.020900461822748184
step: 300, loss: 0.00034734749351628125
step: 310, loss: 0.008450799621641636
step: 320, loss: 0.011712003499269485
step: 330, loss: 0.004073506686836481
step: 340, loss: 0.005216115154325962
step: 350, loss: 0.10848254710435867
step: 360, loss: 0.005304430145770311
step: 370, loss: 0.007325765211135149
step: 380, loss: 0.012040993198752403
step: 390, loss: 0.007915139198303223
step: 400, loss: 0.007978062145411968
step: 410, loss: 0.004889741074293852
step: 420, loss: 0.012387010268867016
step: 430, loss: 0.022842280566692352
step: 440, loss: 0.01878936216235161
step: 450, loss: 0.009008827619254589
step: 460, loss: 0.0010547420242801309
step: 470, loss: 0.001226041349582374
step: 480, loss: 0.003348099533468485
step: 490, loss: 0.018589777871966362
step: 500, loss: 0.017807744443416595
step: 510, loss: 0.00707599613815546
step: 520, loss: 0.06639277935028076
step: 530, loss: 0.0013409419916570187
/usr/local/lib/python3.7/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=0.9252336448598131, f1=0.9208430913348946,
best_f1=0.9262295081967213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.003020698670297861
step: 10, loss: 0.013742104172706604
step: 20, loss: 0.00040516507579013705
step: 30, loss: 0.0011184628820046782
step: 40, loss: 0.0848817527294159
step: 50, loss: 0.004493743646889925
step: 60, loss: 0.016885368153452873
step: 70, loss: 0.00024508702335879207
step: 80, loss: 0.018488241359591484
step: 90, loss: 0.000709176470991224
step: 100, loss: 0.0005691568367183208
step: 110, loss: 0.006137687247246504
step: 120, loss: 0.0048036049120128155
step: 130, loss: 0.005528314970433712
step: 140, loss: 0.001039354596287012
step: 150, loss: 0.00021347191068343818
step: 160, loss: 0.02322983928024769
step: 170, loss: 0.005996611900627613
step: 180, loss: 0.00036525618634186685
step: 190, loss: 0.005818779580295086
step: 200, loss: 0.06351029872894287
step: 210, loss: 0.0027205843944102526
step: 220, loss: 0.0010232918430119753
step: 230, loss: 0.0009209811105392873
step: 240, loss: 0.0029336188454180956
step: 250, loss: 0.002495105378329754
step: 260, loss: 0.02570405602455139
step: 270, loss: 0.061670511960983276
step: 280, loss: 0.04292025417089462
step: 290, loss: 0.001229440444149077
step: 300, loss: 0.04158122465014458
step: 310, loss: 0.001904283999465406
step: 320, loss: 0.0033454718068242073
step: 330, loss: 0.0012260065414011478
step: 340, loss: 0.002236570231616497
step: 350, loss: 0.03400257229804993
step: 360, loss: 0.09639772772789001
step: 370, loss: 0.004312851931899786
step: 380, loss: 0.044617462903261185
step: 390, loss: 0.005213762167841196
step: 400, loss: 0.014500747434794903
step: 410, loss: 0.003424867056310177
step: 420, loss: 0.0028941447380930185
step: 430, loss: 0.003233799012377858
step: 440, loss: 0.030549511313438416
step: 450, loss: 0.031145961955189705
step: 460, loss: 0.011486707255244255
step: 470, loss: 0.014413091354072094
```

```
step: 480, loss: 0.0008992048678919673
step: 490, loss: 0.000668585067614913
step: 500, loss: 0.0004283348098397255
step: 510, loss: 0.000409600354032591
step: 520, loss: 0.0009851471986621618
step: 530, loss: 0.0314522348344326
/usr/local/lib/python3.7/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=0.9174055062995801, f1=0.9212007504690432,
best_f1=0.9262295081967213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.043913740664720535
step: 10, loss: 0.0021557966247200966
step: 20, loss: 0.0004955307231284678
step: 30, loss: 0.00716111995279789
step: 40, loss: 0.0008821478695608675
step: 50, loss: 0.0018971188692376018
step: 60, loss: 0.00013426662189885974
step: 70, loss: 0.027284028008580208
step: 80, loss: 0.0004222407878842205
step: 90, loss: 0.0001885388046503067
step: 100, loss: 0.0007173041230998933
step: 110, loss: 0.016390269622206688
step: 120, loss: 0.0010643585119396448
step: 130, loss: 0.0015764483250677586
step: 140, loss: 0.0011209659278392792
step: 150, loss: 0.0006296908832155168
step: 160, loss: 8.676892321091145e-05
step: 170, loss: 0.0003626719117164612
step: 180, loss: 9.785064321476966e-05
step: 190, loss: 0.003213676158338785
step: 200, loss: 0.0004145743732806295
step: 210, loss: 0.0006517974543385208
step: 220, loss: 0.0002571612421888858
step: 230, loss: 0.0062878746539354324
step: 240, loss: 0.00241635600104928
step: 250, loss: 0.0015631378628313541
step: 260, loss: 0.0019085052190348506
step: 270, loss: 0.0006067389622330666
step: 280, loss: 0.018721316009759903
step: 290, loss: 0.00017022299289237708
step: 300, loss: 0.007325523067265749
step: 310, loss: 0.00034587905975058675
step: 320, loss: 0.0006756268558092415
step: 330, loss: 0.0009888691129162908
```

```
step: 340, loss: 0.0004060055362060666
step: 350, loss: 0.02222432568669319
step: 360, loss: 0.0003278757503721863
step: 370, loss: 0.007488792296499014
step: 380, loss: 0.003223469713702798
step: 390, loss: 0.0011890711029991508
step: 400, loss: 0.00010444447252666578
step: 410, loss: 0.00017434671462979168
step: 420, loss: 0.002170354127883911
step: 430, loss: 0.02757525071501732
step: 440, loss: 0.02156040444970131
step: 450, loss: 0.0004830807156395167
step: 460, loss: 0.02085840329527855
step: 470, loss: 0.06424272805452347
step: 480, loss: 0.0031897802837193012
step: 490, loss: 0.0001813939306885004
step: 500, loss: 0.0007539921789430082
step: 510, loss: 0.00011378966155461967
step: 520, loss: 0.0022557105403393507
step: 530, loss: 0.03683088719844818
/usr/local/lib/python3.7/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=0.9192431933548684, f1=0.9205328433624255,
best_f1=0.9262295081967213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021075403783470392
step: 10, loss: 0.05146021768450737
step: 20, loss: 0.0022198352962732315
step: 30, loss: 0.0019152595195919275
step: 40, loss: 0.0001880962518043816
step: 50, loss: 0.001846402301453054
step: 60, loss: 0.0037904370110481977
step: 70, loss: 0.1244802474975586
step: 80, loss: 0.009494180791079998
step: 90, loss: 0.0001589318853802979
step: 100, loss: 0.0033256174065172672
step: 110, loss: 0.0004173285560682416
step: 120, loss: 0.00014657863357570022
step: 130, loss: 0.00033651621197350323
step: 140, loss: 0.0032288245856761932
step: 150, loss: 0.00015938088472466916
step: 160, loss: 0.004939064383506775
step: 170, loss: 0.0001224530569743365
step: 180, loss: 7.473405275959522e-05
step: 190, loss: 0.000105490064015612
```

```
step: 200, loss: 0.005399594083428383
step: 210, loss: 0.055948324501514435
step: 220, loss: 0.00012947682989761233
step: 230, loss: 0.00021610916883219033
step: 240, loss: 0.005126208066940308
step: 250, loss: 0.0001365221687592566
step: 260, loss: 0.0023499473463743925
step: 270, loss: 6.186384416650981e-05
step: 280, loss: 0.0039454055950045586
step: 290, loss: 0.002320874249562621
step: 300, loss: 9.728854638524354e-05
step: 310, loss: 0.00016477380995638669
step: 320, loss: 0.02314520627260208
step: 330, loss: 9.66447769314982e-05
step: 340, loss: 0.00016344837786164135
step: 350, loss: 0.003066327655687928
step: 360, loss: 0.0023047991562634706
step: 370, loss: 9.846824104897678e-05
step: 380, loss: 0.0007341621676459908
step: 390, loss: 3.989655670011416e-05
step: 400, loss: 5.764820525655523e-05
step: 410, loss: 0.00541083374992013
step: 420, loss: 0.0007925801910459995
step: 430, loss: 0.000421917560743168
step: 440, loss: 8.694636926520616e-05
step: 450, loss: 0.000620267353951931
step: 460, loss: 0.0006600447813980281
step: 470, loss: 0.01959136314690113
step: 480, loss: 0.0044736373238265514
step: 490, loss: 0.00013667884923052043
step: 500, loss: 0.002267146483063698
step: 510, loss: 0.0010275780223309994
step: 520, loss: 0.001026707235723734
step: 530, loss: 0.0015121380565688014
/usr/local/lib/python3.7/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=0.9184834123222749, f1=0.9132075471698115,
best_f1=0.9262295081967213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022719597909599543
step: 10, loss: 0.0019207628211006522
step: 20, loss: 0.03708187863230705
step: 30, loss: 0.01162948552519083
step: 40, loss: 0.0025822760071605444
step: 50, loss: 0.03064783662557602
```

```
step: 60, loss: 0.00024002010468393564
step: 70, loss: 0.00011974003427894786
step: 80, loss: 4.515997716225684e-05
step: 90, loss: 0.0009954518172889948
step: 100, loss: 0.0032478137873113155
step: 110, loss: 0.00046173849841579795
step: 120, loss: 0.002222607145085931
step: 130, loss: 0.011669748462736607
step: 140, loss: 3.2122228731168434e-05
step: 150, loss: 0.00045896272058598697
step: 160, loss: 0.00032347877277061343
step: 170, loss: 0.005521866958588362
step: 180, loss: 0.0005183101748116314
step: 190, loss: 0.0007253431249409914
step: 200, loss: 0.0001483049854869023
step: 210, loss: 0.0004109102301299572
step: 220, loss: 0.0010391032556071877
step: 230, loss: 0.001015104353427887
step: 240, loss: 0.0001970917801372707
step: 250, loss: 0.0009561093174852431
step: 260, loss: 0.008645176887512207
step: 270, loss: 7.605589780723676e-05
step: 280, loss: 0.03796406462788582
step: 290, loss: 0.0015682110097259283
step: 300, loss: 0.000575049314647913
step: 310, loss: 0.011084852740168571
step: 320, loss: 0.0009163098875433207
step: 330, loss: 0.025614852085709572
step: 340, loss: 0.0014062337577342987
step: 350, loss: 0.0007502884254790843
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 360, loss: 0.0003573525173123926
step: 370, loss: 0.0011972181964665651
step: 380, loss: 0.00024994046543724835
step: 390, loss: 0.0001995430066017434
step: 400, loss: 0.028896840289235115
step: 410, loss: 0.00012494805559981614
step: 420, loss: 0.002421749522909522
step: 430, loss: 0.0001881937641883269
step: 440, loss: 0.00042119590216316283
step: 450, loss: 0.008435266092419624
step: 460, loss: 0.003394808154553175
step: 470, loss: 0.0016237737145274878
step: 480, loss: 0.0003929326485376805
step: 490, loss: 0.0001477040641475469
step: 500, loss: 0.0018702055094763637
step: 510, loss: 0.0013254642253741622
step: 520, loss: 0.00019348286150489002
```

```
step: 530, loss: 0.00015367576270364225
/usr/local/lib/python3.7/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=0.8955223880597015, f1=0.8865775136206042,
best_f1=0.9262295081967213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006081154569983482
step: 10, loss: 0.01601432077586651
step: 20, loss: 0.00020037601643707603
step: 30, loss: 0.0004258671833667904
step: 40, loss: 0.006048772018402815
step: 50, loss: 0.0012296403292566538
step: 60, loss: 0.00021091560483910143
step: 70, loss: 0.10482281446456909
step: 80, loss: 0.0020987032912671566
step: 90, loss: 0.0024977826979011297
step: 100, loss: 0.00013677474635187536
step: 110, loss: 0.0034674908965826035
step: 120, loss: 7.930553692858666e-05
step: 130, loss: 0.011996312066912651
step: 140, loss: 0.047384291887283325
step: 150, loss: 0.0001633983920328319
step: 160, loss: 0.0006122516351751983
step: 170, loss: 0.0019939097110182047
step: 180, loss: 4.971679300069809e-05
step: 190, loss: 0.000322941254125908
step: 200, loss: 0.0034482446499168873
step: 210, loss: 5.6483586376998574e-05
step: 220, loss: 6.458389543695375e-05
step: 230, loss: 0.00010782815661514178
step: 240, loss: 0.0005603316240012646
step: 250, loss: 0.0003388807235751301
step: 260, loss: 0.04432421922683716
step: 270, loss: 0.0004860180779360235
step: 280, loss: 9.838151891017333e-05
step: 290, loss: 6.367971946019679e-05
step: 300, loss: 0.0019969248678535223
step: 310, loss: 3.738501982297748e-05
step: 320, loss: 7.970556180225685e-05
step: 330, loss: 0.11262410134077072
step: 340, loss: 0.0006152126588858664
step: 350, loss: 0.00042555268737487495
step: 360, loss: 0.010305630043148994
step: 370, loss: 0.00202431483194232
step: 380, loss: 4.949224239680916e-05
```

```
step: 390, loss: 0.00017632130766287446
step: 400, loss: 0.0011691511608660221
step: 410, loss: 6.921312888152897e-05
step: 420, loss: 3.3593638363527134e-05
step: 430, loss: 3.310558895464055e-05
step: 440, loss: 0.007242511957883835
step: 450, loss: 5.439450978883542e-05
step: 460, loss: 0.00011782138608396053
step: 470, loss: 0.0019334278767928481
step: 480, loss: 9.239991777576506e-05
step: 490, loss: 0.005457546096295118
step: 500, loss: 0.0010477559408172965
step: 510, loss: 0.0004200291878078133
step: 520, loss: 4.0177939808927476e-05
step: 530, loss: 0.00021797780937049538
/usr/local/lib/python3.7/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=0.9175355450236966, f1=0.9111214518380641,
best_f1=0.9262295081967213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03857511281967163
step: 10, loss: 0.00023282840265892446
step: 20, loss: 0.000448839069576934
step: 30, loss: 0.00015248710406012833
step: 40, loss: 0.00011074253416154534
step: 50, loss: 0.00023155465896707028
step: 60, loss: 8.569068450015038e-05
step: 70, loss: 0.0011710546677932143
step: 80, loss: 0.0009768325835466385
step: 90, loss: 9.61652840487659e-05
step: 100, loss: 0.00013065809616819024
step: 110, loss: 0.0015608287649229169
step: 120, loss: 7.532074232585728e-05
step: 130, loss: 5.6769615184748545e-05
step: 140, loss: 5.084294389234856e-05
step: 150, loss: 8.494479698128998e-05
step: 160, loss: 0.0006656579207628965
step: 170, loss: 0.0024473569355905056
step: 180, loss: 0.0010558412177488208
step: 190, loss: 0.00017319212201982737
step: 200, loss: 5.634826447931118e-05
step: 210, loss: 6.0083351854700595e-05
step: 220, loss: 3.126532828900963e-05
step: 230, loss: 0.00014209302025847137
step: 240, loss: 3.504992855596356e-05
```

```
step: 250, loss: 0.00016689891344867647
step: 260, loss: 9.255000622943044e-05
step: 270, loss: 4.7532073949696496e-05
step: 280, loss: 4.0838192944647744e-05
step: 290, loss: 4.217595051159151e-05
step: 300, loss: 0.00010398220911156386
step: 310, loss: 0.00020332785788923502
step: 320, loss: 2.9726748834946193e-05
step: 330, loss: 0.00024261385260615498
step: 340, loss: 5.842879181727767e-05
step: 350, loss: 0.00040122587233781815
step: 360, loss: 4.989911030861549e-05
step: 370, loss: 0.001295228023082018
step: 380, loss: 6.641672371188179e-05
step: 390, loss: 3.6978963180445135e-05
step: 400, loss: 0.0011256623547524214
step: 410, loss: 5.637450885842554e-05
step: 420, loss: 4.001505294581875e-05
step: 430, loss: 4.48531391157303e-05
step: 440, loss: 6.60508667351678e-05
step: 450, loss: 0.0014375792816281319
step: 460, loss: 0.00012202109792269766
step: 470, loss: 2.5674164135125466e-05
step: 480, loss: 0.007199662271887064
step: 490, loss: 0.023641468957066536
step: 500, loss: 0.08841095119714737
step: 510, loss: 0.007000803016126156
step: 520, loss: 5.414545375970192e-05
step: 530, loss: 2.583404966571834e-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=0.9236533957845432, f1=0.9215777262180975,
best_f1=0.9262295081967213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015702260425314307
step: 10, loss: 0.0013030291302129626
step: 20, loss: 3.759482206078246e-05
step: 30, loss: 9.038283315021545e-05
step: 40, loss: 0.0013939959462732077
step: 50, loss: 7.96995300333947e-05
step: 60, loss: 6.337374361464754e-05
step: 70, loss: 6.15192620898597e-05
step: 80, loss: 0.00022186411661095917
step: 90, loss: 8.634741971036419e-05
step: 100, loss: 0.00012473415699787438
```

```
step: 110, loss: 8.526913006789982e-05
step: 120, loss: 0.0004908525734208524
step: 130, loss: 3.5710800148081034e-05
step: 140, loss: 2.8586840926436707e-05
step: 150, loss: 7.192615157691762e-05
step: 160, loss: 0.0003961001639254391
step: 170, loss: 0.00023500782845076174
step: 180, loss: 5.9174511989112943e-05
step: 190, loss: 4.583698319038376e-05
step: 200, loss: 0.0006492853863164783
step: 210, loss: 0.00021500761795323342
step: 220, loss: 0.0001504337415099144
step: 230, loss: 0.00905066728591919
step: 240, loss: 3.574554284568876e-05
step: 250, loss: 9.858019620878622e-05
step: 260, loss: 2.74246067419881e-05
step: 270, loss: 0.0005568523774854839
step: 280, loss: 0.0008206338970921934
step: 290, loss: 0.007170137483626604
step: 300, loss: 0.0012546175858005881
step: 310, loss: 0.002818997949361801
step: 320, loss: 0.011035870760679245
step: 330, loss: 0.0009328182204626501
step: 340, loss: 2.56140592682641e-05
step: 350, loss: 0.0009282805840484798
step: 360, loss: 0.00014810335414949805
step: 370, loss: 0.004213124047964811
step: 380, loss: 2.213893094449304e-05
step: 390, loss: 0.003996847663074732
step: 400, loss: 2.6001527658081613e-05
step: 410, loss: 0.000121555473015178
step: 420, loss: 0.00017810659483075142
step: 430, loss: 4.0662103856448084e-05
step: 440, loss: 8.379895007237792e-05
step: 450, loss: 0.00015348676242865622
step: 460, loss: 0.0009566874941810966
step: 470, loss: 0.0020184810273349285
step: 480, loss: 3.561257472028956e-05
step: 490, loss: 0.01048880722373724
step: 500, loss: 3.864820973831229e-05
step: 510, loss: 2.8449116143747233e-05
step: 520, loss: 4.837457890971564e-05
step: 530, loss: 6.651003786828369e-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=0.927170868347339, f1=0.9225776541492814,
best_f1=0.9262295081967213
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004702146630734205
step: 10, loss: 4.183920827927068e-05
step: 20, loss: 2.7606987714534625e-05
step: 30, loss: 3.615995228756219e-05
step: 40, loss: 3.98763149860315e-05
step: 50, loss: 0.0013497304171323776
step: 60, loss: 0.003472050651907921
step: 70, loss: 0.0013233944773674011
step: 80, loss: 0.0004672174109145999
step: 90, loss: 0.00018132709374185652
step: 100, loss: 0.006409997586160898
step: 110, loss: 2.063403371721506e-05
step: 120, loss: 5.525954838958569e-05
step: 130, loss: 0.00011272636766079813
step: 140, loss: 3.737341830856167e-05
step: 150, loss: 0.001900423550978303
step: 160, loss: 4.64498189103324e-05
step: 170, loss: 2.331986979697831e-05
step: 180, loss: 0.0013287916081026196
step: 190, loss: 3.520845712046139e-05
step: 200, loss: 0.00011784688831539825
step: 210, loss: 0.00016462586063425988
step: 220, loss: 4.522268500295468e-05
step: 230, loss: 8.672229159856215e-05
step: 240, loss: 0.00016769573267083615
step: 250, loss: 3.3872725907713175e-05
step: 260, loss: 8.286933734780177e-05
step: 270, loss: 4.527446799329482e-05
step: 280, loss: 6.351856427500024e-05
step: 290, loss: 7.890521374065429e-05
step: 300, loss: 0.00022249481116887182
step: 310, loss: 4.6603756345575675e-05
step: 320, loss: 9.452475933358073e-05
step: 330, loss: 2.837477222783491e-05
step: 340, loss: 6.683800165774301e-05
step: 350, loss: 0.03150178864598274
step: 360, loss: 0.0001497122284490615
step: 370, loss: 0.00019288594194222242
step: 380, loss: 3.326938531245105e-05
step: 390, loss: 3.440431464696303e-05
step: 400, loss: 3.69670451618731e-05
step: 410, loss: 5.239816164248623e-05
step: 420, loss: 4.941301085636951e-05
step: 430, loss: 3.461267988313921e-05
step: 440, loss: 0.00014232657849788666
```

```
step: 450, loss: 0.00035069958539679646
step: 460, loss: 6.598985055461526e-05
step: 470, loss: 3.4583110391395167e-05
step: 480, loss: 4.023102155770175e-05
step: 490, loss: 5.322071228874847e-05
step: 500, loss: 0.0005202386528253555
step: 510, loss: 5.3542691603070125e-05
step: 520, loss: 0.012106127105653286
step: 530, loss: 4.386913133203052e-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=0.9251059821008009, f1=0.9129213483146068,
best_f1=0.9262295081967213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017977379320655018
step: 10, loss: 3.486876448732801e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.021427195519208908
step: 30, loss: 0.0009799464605748653
step: 40, loss: 2.1159254174563102e-05
step: 50, loss: 8.949147741077468e-05
step: 60, loss: 3.155926062268622e-05
step: 70, loss: 2.7863796276506037e-05
step: 80, loss: 4.731428634840995e-05
step: 90, loss: 0.0003538171877153218
step: 100, loss: 2.1717933122999966e-05
step: 110, loss: 4.512131636147387e-05
step: 120, loss: 3.8959715311648324e-05
step: 130, loss: 3.919897790183313e-05
step: 140, loss: 3.746816219063476e-05
step: 150, loss: 2.664236853888724e-05
step: 160, loss: 4.6428795030806214e-05
step: 170, loss: 2.4396336812060326e-05
step: 180, loss: 3.7454516132129356e-05
step: 190, loss: 3.604077210184187e-05
step: 200, loss: 2.1699212084058672e-05
step: 210, loss: 0.0016910855192691088
step: 220, loss: 2.466055411787238e-05
step: 230, loss: 1.4241650205804035e-05
step: 240, loss: 5.0070466386387125e-05
step: 250, loss: 0.03262210637331009
step: 260, loss: 6.334891804726794e-05
step: 270, loss: 1.9602150132413954e-05
step: 280, loss: 7.561760867247358e-05
step: 290, loss: 0.00046627316623926163
```

```
step: 300, loss: 8.615779370302334e-05
step: 310, loss: 2.7367836082703434e-05
step: 320, loss: 6.335194484563544e-05
step: 330, loss: 6.945734639884904e-05
step: 340, loss: 3.3280222851317376e-05
step: 350, loss: 0.0005618603900074959
step: 360, loss: 4.625396468327381e-05
step: 370, loss: 4.860651824856177e-05
step: 380, loss: 0.00012800187687389553
step: 390, loss: 6.213612505234778e-05
step: 400, loss: 2.1393609131337143e-05
step: 410, loss: 2.66275746980682e-05
step: 420, loss: 3.5495995689416304e-05
step: 430, loss: 0.00047715642722323537
step: 440, loss: 6.69760411255993e-05
step: 450, loss: 6.200626376084983e-05
step: 460, loss: 2.2105024982010946e-05
step: 470, loss: 0.00018318115326110274
step: 480, loss: 3.0098466595518403e-05
step: 490, loss: 5.840997619088739e-05
step: 500, loss: 8.490416803397238e-05
step: 510, loss: 6.404077430488542e-05
step: 520, loss: 0.0013035431038588285
step: 530, loss: 2.874328856705688e-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=0.9239489844119037, f1=0.9159229685298264,
best_f1=0.9262295081967213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.961210263194516e-05
step: 10, loss: 0.0015011890791356564
step: 20, loss: 0.0005819791695103049
step: 30, loss: 6.149283581180498e-05
step: 40, loss: 0.00010766583000076935
step: 50, loss: 1.9728791812667623e-05
step: 60, loss: 2.4254710297100246e-05
step: 70, loss: 1.978840009542182e-05
step: 80, loss: 0.000259447522694245
step: 90, loss: 0.0001724495377857238
step: 100, loss: 3.8241683796513826e-05
step: 110, loss: 2.6027291823993437e-05
step: 120, loss: 1.575776877871249e-05
step: 130, loss: 1.8160506442654878e-05
step: 140, loss: 0.0009674854227341712
step: 150, loss: 2.569261414464563e-05
```

```
step: 160, loss: 2.6336547307437286e-05
step: 170, loss: 3.2002877560444176e-05
step: 180, loss: 0.04937507584691048
step: 190, loss: 2.9156650271033868e-05
step: 200, loss: 3.266942803747952e-05
step: 210, loss: 0.0014299947069957852
step: 220, loss: 2.6526759029366076e-05
step: 230, loss: 2.731278436840512e-05
step: 240, loss: 2.3658609279664233e-05
step: 250, loss: 5.807128763990477e-05
step: 260, loss: 0.0003819196135737002
step: 270, loss: 0.0001240877027157694
step: 280, loss: 1.5329411326092668e-05
step: 290, loss: 2.5442599508096464e-05
step: 300, loss: 2.2038124370737933e-05
step: 310, loss: 1.8503173123463057e-05
step: 320, loss: 0.003504939377307892
step: 330, loss: 0.00011981092393398285
step: 340, loss: 2.8069393010810018e-05
step: 350, loss: 0.00018000898126047105
step: 360, loss: 2.3200582290883176e-05
step: 370, loss: 2.281669912918005e-05
step: 380, loss: 0.00010404809290776029
step: 390, loss: 4.136160350753926e-05
step: 400, loss: 3.054260741919279e-05
step: 410, loss: 1.9415898350416683e-05
step: 420, loss: 1.6610834791208617e-05
step: 430, loss: 4.965777407051064e-05
step: 440, loss: 1.6357495042029768e-05
step: 450, loss: 1.7311194824287668e-05
step: 460, loss: 5.3158575610723346e-05
step: 470, loss: 1.9765911929425783e-05
step: 480, loss: 3.624411328928545e-05
step: 490, loss: 2.0522207705653273e-05
step: 500, loss: 1.911787512653973e-05
step: 510, loss: 3.0720180802745745e-05
step: 520, loss: 2.6001862352131866e-05
step: 530, loss: 8.836860070005059e-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=0.9249193919852603, f1=0.9237790963030579,
best_f1=0.9262295081967213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005127934855408967
step: 10, loss: 4.777143112733029e-05
```

```
step: 20, loss: 3.08736962324474e-05
step: 30, loss: 2.6657473426894285e-05
step: 40, loss: 1.926671393448487e-05
step: 50, loss: 0.00020178541308268905
step: 60, loss: 7.790025847498327e-05
step: 70, loss: 3.620416100602597e-05
step: 80, loss: 2.876544022001326e-05
step: 90, loss: 4.873410580330528e-05
step: 100, loss: 2.4112900064210407e-05
step: 110, loss: 4.093664028914645e-05
step: 120, loss: 5.1963404985144734e-05
step: 130, loss: 0.00012681307271122932
step: 140, loss: 1.7322341591352597e-05
step: 150, loss: 2.101019344991073e-05
step: 160, loss: 6.950399983907118e-05
step: 170, loss: 3.812032809946686e-05
step: 180, loss: 2.1848163669346832e-05
step: 190, loss: 5.0238224503118545e-05
step: 200, loss: 2.1516831111512147e-05
step: 210, loss: 5.160293585504405e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 220, loss: 3.178766928613186e-05
step: 230, loss: 2.2395737687475048e-05
step: 240, loss: 0.0003900878655258566
step: 250, loss: 8.701834303792566e-05
step: 260, loss: 2.4288281565532088e-05
step: 270, loss: 0.0008871537866070867
step: 280, loss: 3.500462116790004e-05
step: 290, loss: 2.987194966408424e-05
step: 300, loss: 0.001798237208276987
step: 310, loss: 2.4452041543554515e-05
step: 320, loss: 3.063940312131308e-05
step: 330, loss: 0.0007682116120122373
step: 340, loss: 5.7660621678223833e-05
step: 350, loss: 3.346138328197412e-05
step: 360, loss: 2.8754506274708547e-05
step: 370, loss: 1.6372458048863336e-05
step: 380, loss: 2.9342054403969087e-05
step: 390, loss: 2.8540443963720463e-05
step: 400, loss: 8.084779983619228e-05
step: 410, loss: 3.236388511140831e-05
step: 420, loss: 3.7831188819836825e-05
step: 430, loss: 4.50538209406659e-05
step: 440, loss: 0.0001292702800128609
step: 450, loss: 3.873151945299469e-05
step: 460, loss: 7.566428394056857e-05
step: 470, loss: 0.00018079382425639778
step: 480, loss: 1.8350470782024786e-05
```

```
step: 490, loss: 0.002256155014038086
step: 500, loss: 5.82004759053234e-05
step: 510, loss: 4.5417375076795e-05
step: 520, loss: 5.539022822631523e-05
step: 530, loss: 2.3625274479854852e-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=0.9235127478753541, f1=0.9163533834586467,
best_f1=0.9262295081967213
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.")
5742it [00:16, 338.64it/s]
load_f1 = 0.9280373831775701
real_f1 = 0.9263059701492538
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 [00:12, 346.92it/s]
```

3 DK

```
# Para executar 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 1.1 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: 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: 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: 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: 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: 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: 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: 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: 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: 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: 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: 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: 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: 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: 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=a015ae3847c2205d04140e7ceb414541ad25b990e9c341b9461d751226e12dc7
  Stored in directory: /tmp/pip-ephem-wheel-cache-5mg0132o/wheels/11/95/ba/2c36c
c368c0bd339b44a791c2c1881a1fb714b78c29a4cb8f5
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/
```

```
Downloading: 100% 442/442 [00:00<00:00, 504kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 2.59MB/s]
Downloading: 100% 268M/268M [00:03<00:00, 70.5MB/s]
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: 0.8580923676490784
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=0.39999999999999997, 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: 0.3682679235935211
/usr/local/lib/python3.7/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=0.3783783783783784, f1=0.3404255319148936,
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: 0.359671950340271
/usr/local/lib/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=0.5185185185185186, f1=0.3243243243243243,
best_f1=0.3243243243243243
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3618074357509613
/usr/local/lib/python3.7/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=0.5128205128205129, f1=0.3404255319148936,
best_f1=0.3243243243243243
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.32989615201950073
/usr/local/lib/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=0.5909090909090909, f1=0.36363636363636365,
best_f1=0.36363636363636365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2963029444217682
/usr/local/lib/python3.7/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=0.5581395348837208, f1=0.40816326530612246,
best_f1=0.36363636363636365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2332436740398407
/usr/local/lib/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=0.631578947368421, f1=0.38095238095238093,
best_f1=0.38095238095238093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.32686421275138855
/usr/local/lib/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=0.7272727272727273, f1=0.43750000000000006,
best_f1=0.43750000000000006
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18289241194725037
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 9: dev_f1=0.7878787878787878, f1=0.43750000000000006,
best_f1=0.43750000000000006
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19856742024421692
/usr/local/lib/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=0.8148148148148148, f1=0.4799999999999999,
best_f1=0.4799999999999999
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23542624711990356
/usr/local/lib/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=0.8275862068965518, f1=0.4615384615384615,
best_f1=0.4615384615384615
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14015860855579376
/usr/local/lib/python3.7/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=0.8148148148148148, f1=0.42857142857142855,
best_f1=0.4615384615384615
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09630250185728073
/usr/local/lib/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=0.8571428571428571, f1=0.4666666666666667,
best_f1=0.4666666666666667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18377640843391418
/usr/local/lib/python3.7/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=0.8461538461538461, f1=0.3846153846153846,
best_f1=0.4666666666666667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1329357624053955
```

```
/usr/local/lib/python3.7/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=0.8461538461538461, f1=0.3846153846153846,
best_f1=0.4666666666666667
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, 124895.83it/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 = 0.6086956521739131
real_f1 = 0.6666666666666666
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, 440.54it/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 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: 0.7998037934303284
step: 10, loss: 0.48122069239616394
step: 20, loss: 0.5860881209373474
step: 30, loss: 0.4173172414302826
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 40, loss: 0.13621704280376434
step: 50, loss: 0.06606712192296982
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.16773156821727753
step: 70, loss: 0.21423524618148804
step: 80, loss: 0.1871095597743988
step: 90, loss: 0.034771066159009933
step: 100, loss: 0.15412624180316925
step: 110, loss: 0.1093568503856659
step: 120, loss: 0.013984029181301594
step: 130, loss: 0.017564591020345688
step: 140, loss: 0.031160868704319
step: 150, loss: 0.09042870998382568
step: 160, loss: 0.1492447555065155
step: 170, loss: 0.004413170740008354
step: 180, loss: 0.014233464375138283
step: 190, loss: 0.10223749279975891
step: 200, loss: 0.0072797974571585655
step: 210, loss: 0.007450975943356752
step: 220, loss: 0.008479481562972069
```

```
step: 230, loss: 0.038011688739061356
/usr/local/lib/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=0.9722530521642618, f1=0.9647326507394768,
best_f1=0.9647326507394768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1150221973657608
step: 10, loss: 0.0015200419584289193
step: 20, loss: 0.005840816535055637
step: 30, loss: 0.03375997394323349
step: 40, loss: 0.00991396140307188
step: 50, loss: 0.14034654200077057
step: 60, loss: 0.004484008066356182
step: 70, loss: 0.07539843767881393
step: 80, loss: 0.003757874481379986
step: 90, loss: 0.006823844276368618
step: 100, loss: 0.10402321815490723
step: 110, loss: 0.09161697328090668
step: 120, loss: 0.010243009775876999
step: 130, loss: 0.03413502871990204
step: 140, loss: 0.0763528048992157
step: 150, loss: 0.008649151772260666
step: 160, loss: 0.010303587652742863
step: 170, loss: 0.04500158503651619
step: 180, loss: 0.0048189531080424786
step: 190, loss: 0.2892153263092041
step: 200, loss: 0.013651369139552116
step: 210, loss: 0.049951475113630295
step: 220, loss: 0.0042184037156403065
step: 230, loss: 0.04262998700141907
/usr/local/lib/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=0.9753363228699552, f1=0.9819413092550789,
best_f1=0.9819413092550789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10013590008020401
step: 10, loss: 0.012151917442679405
step: 20, loss: 0.03834014758467674
step: 30, loss: 0.027593137696385384
step: 40, loss: 0.008090265095233917
step: 50, loss: 0.0030586551874876022
step: 60, loss: 0.0034924058709293604
step: 70, loss: 0.001833889284171164
step: 80, loss: 0.0015340055106207728
step: 90, loss: 0.0012456001713871956
step: 100, loss: 0.018744712695479393
step: 110, loss: 0.07628022879362106
step: 120, loss: 0.0038172882050275803
step: 130, loss: 0.03168611228466034
step: 140, loss: 0.003188511822372675
step: 150, loss: 0.00204146234318614
step: 160, loss: 0.0033050307538360357
step: 170, loss: 0.006056994665414095
step: 180, loss: 0.003467816859483719
step: 190, loss: 0.08135545998811722
step: 200, loss: 0.0012300063390284777
step: 210, loss: 0.09782440215349197
step: 220, loss: 0.0017706670332700014
step: 230, loss: 0.03986199572682381
/usr/local/lib/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=0.9800443458980044, f1=0.977728285077951,
best_f1=0.977728285077951
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002249338198453188
step: 10, loss: 0.002369492081925273
step: 20, loss: 0.0018587848171591759
step: 30, loss: 0.00030877121025696397
step: 40, loss: 0.004616524558514357
step: 50, loss: 0.0016325502656400204
step: 60, loss: 0.0007575160707347095
step: 70, loss: 0.01819567009806633
step: 80, loss: 0.0056762052699923515
step: 90, loss: 0.001539053162559867
step: 100, loss: 0.005007600411772728
```

```
step: 110, loss: 0.006859447807073593
step: 120, loss: 0.0036988251376897097
step: 130, loss: 0.0016015386208891869
step: 140, loss: 0.0012824226869270205
step: 150, loss: 0.001500228769145906
step: 160, loss: 0.022867197170853615
step: 170, loss: 0.0030894260853528976
step: 180, loss: 0.09598267823457718
step: 190, loss: 0.005477687809616327
step: 200, loss: 0.0012756349751725793
step: 210, loss: 0.00028197094798088074
step: 220, loss: 0.00092996348394081
step: 230, loss: 0.0002342781631738096
/usr/local/lib/python3.7/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=0.9798657718120806, f1=0.9831271091113611,
best_f1=0.977728285077951
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00023127024178393185
step: 10, loss: 0.0005567007465288043
step: 20, loss: 0.00032921231468208134
step: 30, loss: 0.00015521592285949737
step: 40, loss: 0.0011397902853786945
step: 50, loss: 0.0002584901812952012
step: 60, loss: 0.0008698515011928976
step: 70, loss: 0.0035073254257440567
step: 80, loss: 0.00140906358137726 78
step: 90, loss: 0.012468096800148487
step: 100, loss: 0.06021694093942642
step: 110, loss: 0.002649930538609624
step: 120, loss: 0.03955889493227005
step: 130, loss: 0.142506405711174
step: 140, loss: 0.0008686163928359747
step: 150, loss: 0.0007799084414727986
step: 160, loss: 0.06500299274921417
step: 170, loss: 0.03002045676112175
step: 180, loss: 0.0008010546443983912
step: 190, loss: 0.003356191096827388
step: 200, loss: 0.0070465337485075
step: 210, loss: 0.0004833849670831114
step: 220, loss: 0.0012258688220754266
step: 230, loss: 0.0009333560592494905
/usr/local/lib/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=0.9854423292273236, f1=0.9797297297297298,
best_f1=0.9797297297297298
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009766279254108667
step: 10, loss: 0.0008202052558772266
step: 20, loss: 0.0004959665820933878
step: 30, loss: 0.0006496921996586025
step: 40, loss: 0.0007169743767008185
step: 50, loss: 0.0027756092604249716
step: 60, loss: 0.0011529240291565657
step: 70, loss: 0.0007804962806403637
step: 80, loss: 0.018957240507006645
step: 90, loss: 0.0002935981028713286
step: 100, loss: 0.001333892229013145
step: 110, loss: 0.16226717829704285
step: 120, loss: 0.00033086430630646646
step: 130, loss: 0.0028262895066291094
step: 140, loss: 0.005849407985806465
step: 150, loss: 0.00017343839863315225
step: 160, loss: 0.00041894635069184005
step: 170, loss: 0.0001437133178114891
step: 180, loss: 0.037092048674821854
step: 190, loss: 0.0006412201328203082
step: 200, loss: 0.00020509312162175775
step: 210, loss: 0.0009487827192060649
step: 220, loss: 0.00016525320825167
step: 230, loss: 0.0029200429562479258
/usr/local/lib/python3.7/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=0.9820627802690582, f1=0.9820627802690582,
best_f1=0.9797297297297298
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1379886269569397
step: 10, loss: 0.000380128767574206
step: 20, loss: 0.0002931756025645882
step: 30, loss: 0.0010417981538921595
step: 40, loss: 0.001404638052918017
step: 50, loss: 0.0007038760231807828
step: 60, loss: 0.0006527503719553351
```

```
step: 70, loss: 0.0003892076783813536
step: 80, loss: 0.00012384189176373184
step: 90, loss: 0.00036784756230190396
step: 100, loss: 0.0011725189397111535
step: 110, loss: 0.001846661907620728
step: 120, loss: 0.011445225216448307
step: 130, loss: 0.018578341230750084
step: 140, loss: 9.413827501703054e-05
step: 150, loss: 0.00013974914327263832
step: 160, loss: 9.793826757231727e-05
step: 170, loss: 0.0021101185120642185
step: 180, loss: 0.0005503837601281703
step: 190, loss: 0.0013447034871205688
step: 200, loss: 0.0003806938184425235
step: 210, loss: 0.00012926312047056854
step: 220, loss: 0.0003099520690739155
step: 230, loss: 0.012790961191058159
/usr/local/lib/python3.7/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=0.9831271091113611, f1=0.976324689966178,
best_f1=0.9797297297297298
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01321435160934925
step: 10, loss: 0.0026358391623944044
step: 20, loss: 0.0013224994763731956
step: 30, loss: 0.0003171316930092871
step: 40, loss: 0.00011402348172850907
step: 50, loss: 0.00022376611013896763
step: 60, loss: 0.0006903029279783368
step: 70, loss: 0.008944418281316757
step: 80, loss: 0.003277478041127324
step: 90, loss: 0.00035779085010290146
step: 100, loss: 0.0002758194459602237
step: 110, loss: 0.0005668981466442347
step: 120, loss: 0.00023192734806798398
step: 130, loss: 0.0070975101552903365
step: 140, loss: 0.00010213222412858158
step: 150, loss: 0.00034256192157045007
step: 160, loss: 0.00014455523341894415
step: 170, loss: 0.005143473390489817
step: 180, loss: 0.0010928392875939608
step: 190, loss: 0.0009053087560459971
step: 200, loss: 0.0960891917347908
step: 210, loss: 0.00013488413242157550
step: 220, loss: 0.00011364805686753243
```

```
step: 230, loss: 0.0063599226996302605
/usr/local/lib/python3.7/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=0.9841628959276018, f1=0.9761634506242906,
best_f1=0.9797297297297298
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006135131698101759
step: 10, loss: 0.0002526099851820618
step: 20, loss: 0.0002700758632272482
step: 30, loss: 0.004361102823168039
step: 40, loss: 0.12571385502815247
step: 50, loss: 0.00019503796647768468
step: 60, loss: 0.005896334070712328
step: 70, loss: 0.0004657598037738353
step: 80, loss: 0.020149188116192818
step: 90, loss: 0.0005127073382027447
step: 100, loss: 8.103645086521283e-05
step: 110, loss: 0.001197031233459711
step: 120, loss: 0.0021511418744921684
step: 130, loss: 8.424245606875047e-05
step: 140, loss: 0.0008357961778528988
step: 150, loss: 4.8903824790613726e-05
step: 160, loss: 0.00041681682597845793
step: 170, loss: 6.061475505703129e-05
step: 180, loss: 0.00011916649236809462
step: 190, loss: 0.0005998606793582439
step: 200, loss: 0.0006298673106357455
step: 210, loss: 0.0001343652547802776
step: 220, loss: 0.013424175791442394
step: 230, loss: 0.00021620096231345087
/usr/local/lib/python3.7/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=0.9841628959276018, f1=0.9807474518686297,
best_f1=0.9797297297297298
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002681775949895382
step: 10, loss: 6.851299986010417e-05
step: 20, loss: 0.010785354301333427
step: 30, loss: 0.0004019050975330174
step: 40, loss: 5.48661919310689e-05
step: 50, loss: 0.00036351074231788516
step: 60, loss: 0.0018404594156891108
```

```
step: 70, loss: 9.4034789071884e-05
step: 80, loss: 0.0003482299216557294
step: 90, loss: 0.00018628689576871693
step: 100, loss: 0.00010752672824310139
step: 110, loss: 0.0015154413413256407
step: 120, loss: 0.0009307385771535337
step: 130, loss: 0.001875729882158339
step: 140, loss: 0.0009877549018710852
step: 150, loss: 0.00229831226170063
step: 160, loss: 0.00024604739155620337
step: 170, loss: 8.361427171621472e-05
step: 180, loss: 0.041823532432317734
step: 190, loss: 8.520378469256684e-05
step: 200, loss: 4.920443461742252e-05
step: 210, loss: 0.00013577168283518404
step: 220, loss: 8.747229003347456e-05
step: 230, loss: 7.383626507362351e-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 10: dev_f1=0.9864864864864865, f1=0.9785794813979707,
best_f1=0.9785794813979707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.730894554289989e-05
step: 10, loss: 6.530893733724952e-05
step: 20, loss: 3.765375731745735e-05
step: 30, loss: 0.0004610844189301133
step: 40, loss: 0.0004934114986099303
step: 50, loss: 0.021934542804956436
step: 60, loss: 0.00035806820960715413
step: 70, loss: 7.922131771920249e-05
step: 80, loss: 0.00022127390548121184
step: 90, loss: 0.00033503177110105753
step: 100, loss: 0.0005727841053158045
step: 110, loss: 5.318528201314621e-05
step: 120, loss: 8.723205246496946e-05
step: 130, loss: 6.737458170391619e-05
step: 140, loss: 9.688099817140028e-05
step: 150, loss: 4.1784787754295394e-05
step: 160, loss: 4.699170676758513e-05
step: 170, loss: 0.00040253507904708385
step: 180, loss: 0.0004481365904211998
```

```
step: 190, loss: 0.0001978894870262593
step: 200, loss: 0.0001011622225632891
step: 210, loss: 4.463799996301532e-05
step: 220, loss: 9.711809252621606e-05
step: 230, loss: 0.0005232356488704681
/usr/local/lib/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=0.9909706546275394, 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: 0.00044102902757003903
step: 10, loss: 5.260522084427066e-05
step: 20, loss: 4.27043269155547e-05
step: 30, loss: 7.07023500581272e-05
step: 40, loss: 6.061456224415451e-05
step: 50, loss: 0.0019314860692247748
step: 60, loss: 0.0005012772744521499
step: 70, loss: 9.028964268509299e-05
step: 80, loss: 0.0014192829839885235
step: 90, loss: 0.00010154307528864592
step: 100, loss: 7.287685002665967e-05
step: 110, loss: 0.00012236242764629424
step: 120, loss: 0.00010844458302017301
step: 130, loss: 8.264587086159736e-05
step: 140, loss: 5.7662069593789056e-05
step: 150, loss: 0.00011270376126049086
step: 160, loss: 0.00040457051363773644
step: 170, loss: 0.000131929264171049
step: 180, loss: 0.00010346496128477156
step: 190, loss: 0.0004952350282110274
step: 200, loss: 0.0005662179901264608
step: 210, loss: 0.0005053521599620581
step: 220, loss: 0.0003289137384854257
step: 230, loss: 0.00200834427960217
/usr/local/lib/python3.7/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=0.9767441860465117, f1=0.977728285077951,
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: 0.007771010044962168
step: 10, loss: 0.0023035428021103144
step: 20, loss: 9.959794988390058e-05
step: 30, loss: 0.001128236297518015
step: 40, loss: 0.00012032501399951706
step: 50, loss: 7.094067404977977e-05
step: 60, loss: 0.00014207392814569175
step: 70, loss: 6.832235521869734e-05
step: 80, loss: 0.00019412442634347826
step: 90, loss: 4.620960316970013e-05
step: 100, loss: 6.368728645611554e-05
step: 110, loss: 0.00010278626723447815
step: 120, loss: 8.628680370748043e-05
step: 130, loss: 8.471956971334293e-05
step: 140, loss: 7.182971603469923e-05
step: 150, loss: 0.00014362568617798388
step: 160, loss: 5.512183270184323e-05
step: 170, loss: 5.139108907314949e-05
step: 180, loss: 0.00011980887211393565
step: 190, loss: 4.156818977207877e-05
step: 200, loss: 4.562180765788071e-05
step: 210, loss: 5.996189793222584e-05
step: 220, loss: 4.563130278256722e-05
step: 230, loss: 6.1111080867704e-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=0.9863945578231292, f1=0.9796380090497738,
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: 4.842596536036581e-05
step: 10, loss: 4.0973689465317875e-05
step: 20, loss: 0.000445740093709901
step: 30, loss: 5.897882874705829e-05
step: 40, loss: 4.3488351366249844e-05
step: 50, loss: 6.720793317072093e-05
step: 60, loss: 0.0001285057223867625
step: 70, loss: 0.0024412802886217833
step: 80, loss: 7.008887769188732e-05
step: 90, loss: 8.061707922024652e-05
step: 100, loss: 0.0001005859230645001
step: 110, loss: 0.0002508196630515158
step: 120, loss: 6.346853479044512e-05
step: 130, loss: 7.855745207052678e-05
step: 140, loss: 4.5940429117763415e-05
```

```
step: 150, loss: 6.52472153888084e-05
step: 160, loss: 4.3016985728172585e-05
step: 170, loss: 6.390854105120525e-05
step: 180, loss: 0.00016859252355061471
step: 190, loss: 4.7063502279343084e-05
step: 200, loss: 4.071612420375459e-05
step: 210, loss: 0.012678450904786587
step: 220, loss: 5.633695036522113e-05
step: 230, loss: 7.90442936704494e-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=0.9842342342342343, f1=0.9797297297297298,
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: 5.065152436145581e-05
step: 10, loss: 0.00011774705490097404
step: 20, loss: 8.268449164461344e-05
step: 30, loss: 0.00013097634655423462
step: 40, loss: 0.00010520267824176699
step: 50, loss: 5.959366899332963e-05
step: 60, loss: 0.00010998112702509388
step: 70, loss: 6.401420978363603e-05
step: 80, loss: 0.00020045954443048686
step: 90, loss: 4.083503517904319e-05
step: 100, loss: 0.00011704287317115813
step: 110, loss: 8.315332524944097e-05
step: 120, loss: 6.071360985515639e-05
step: 130, loss: 6.136043521109968e-05
step: 140, loss: 0.0001577392395120114
step: 150, loss: 5.741458880947903e-05
step: 160, loss: 5.069862163509242e-05
step: 170, loss: 4.364259439171292e-05
step: 180, loss: 5.8586738305166364e-05
step: 190, loss: 9.929604857461527e-05
step: 200, loss: 5.765909008914605e-05
step: 210, loss: 9.573289571562782e-05
step: 220, loss: 5.581192090176046e-05
step: 230, loss: 7.870845001889393e-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=0.9842696629213483, f1=0.9797297297297298,
best_f1=0.9807474518686297
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 [00:08, 301.78it/s]
load_f1 = 0.9853107344632768
real_f1 = 0.9841986455981941
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 [00:12, 345.55it/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   : 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: 0.7797459959983826
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.44885021448135376
step: 20, loss: 0.49343833327293396
step: 30, loss: 0.43989869952201843
step: 40, loss: 0.41075509786605835
step: 50, loss: 0.3022240996360779
step: 60, loss: 0.25892284512519836
step: 70, loss: 0.11246855556964874
step: 80, loss: 0.1010415181517601
step: 90, loss: 0.1358979046344757
step: 100, loss: 0.305128812789917
step: 110, loss: 0.1109989807009697
step: 120, loss: 0.028715094551444054
step: 130, loss: 0.022470125928521156
step: 140, loss: 0.19626331329345703
step: 150, loss: 0.028989754617214203
step: 160, loss: 0.10173939913511276
step: 170, loss: 0.24722251296043396
step: 180, loss: 0.09380923956632614
step: 190, loss: 0.020140407606959343
step: 200, loss: 0.08991032093763351
step: 210, loss: 0.07942280918359756
step: 220, loss: 0.07449806481599808
step: 230, loss: 0.06529012322425842
step: 240, loss: 0.06529244035482407
step: 250, loss: 0.07660980522632599
step: 260, loss: 0.04653164744377136
step: 270, loss: 0.015999024733901024
step: 280, loss: 0.13288690149784088
step: 290, loss: 0.08731482923030853
step: 300, loss: 0.05671609938144684
step: 310, loss: 0.09058722108602524
step: 320, loss: 0.0764358639717102
step: 330, loss: 0.0506260022521019
```

```
step: 340, loss: 0.2996246814727783
step: 350, loss: 0.03402046114206314
step: 360, loss: 0.06446975469589233
step: 370, loss: 0.0851183757185936
step: 380, loss: 0.16822457313537598
step: 390, loss: 0.26829633116722107
step: 400, loss: 0.013576898723840714
step: 410, loss: 0.085858054459095
step: 420, loss: 0.03443557024002075
step: 430, loss: 0.07829061895608902
step: 440, loss: 0.2539082467556
step: 450, loss: 0.01245596818625927
step: 460, loss: 0.07447549700737
step: 470, loss: 0.21361781656742096
step: 480, loss: 0.1713077276945114
step: 490, loss: 0.03119860216975212
step: 500, loss: 0.02783745341002941
step: 510, loss: 0.10458528995513916
step: 520, loss: 0.06472288072109222
step: 530, loss: 0.16129116714000702
/usr/local/lib/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=0.9219600725952812, f1=0.9154092363968908,
best_f1=0.9154092363968908
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.132328599691391
step: 10, loss: 0.2137628048658371
step: 20, loss: 0.1639823317527771
step: 30, loss: 0.04720113426446915
step: 40, loss: 0.03897516801953316
step: 50, loss: 0.10706843435764313
step: 60, loss: 0.10974389314651489
step: 70, loss: 0.19961047172546387
step: 80, loss: 0.01729254424571991
step: 90, loss: 0.06619367748498917
step: 100, loss: 0.2156323343515396
step: 110, loss: 0.014207717962563038
step: 120, loss: 0.04947257041931152
step: 130, loss: 0.019126037135720253
step: 140, loss: 0.034969545900821686
step: 150, loss: 0.014398154802620411
```

```
step: 160, loss: 0.03787987679243088
step: 170, loss: 0.09378498047590256
step: 180, loss: 0.005077838432043791
step: 190, loss: 0.037918753921985626
step: 200, loss: 0.0241196621209383
step: 210, loss: 0.025698469951748848
step: 220, loss: 0.10243353992700577
step: 230, loss: 0.03490578010678291
step: 240, loss: 0.10916952788829803
step: 250, loss: 0.01796264760196209
step: 260, loss: 0.01562824659049511
step: 270, loss: 0.18088187277317047
step: 280, loss: 0.21083982288837433
step: 290, loss: 0.1233803778886795
step: 300, loss: 0.0120620122179389
step: 310, loss: 0.10712117701768875
step: 320, loss: 0.15774250030517578
step: 330, loss: 0.08801676332950592
step: 340, loss: 0.10233071446418762
step: 350, loss: 0.0513763390481472
step: 360, loss: 0.07218466699123383
step: 370, loss: 0.05405636504292488
step: 380, loss: 0.06684033572673798
step: 390, loss: 0.02709037810564041
step: 400, loss: 0.04608854278922081
step: 410, loss: 0.0022925958037376404
step: 420, loss: 0.08071637898683548
step: 430, loss: 0.04882750287652016
step: 440, loss: 0.018731895834207535
step: 450, loss: 0.016872864216566086
step: 460, loss: 0.24175670742988586
step: 470, loss: 0.042258668690919876
step: 480, loss: 0.23478227853775024
step: 490, loss: 0.20180104672908783
step: 500, loss: 0.05365228280425072
step: 510, loss: 0.09805913269519806
step: 520, loss: 0.06339596956968307
step: 530, loss: 0.15691915154457092
/usr/local/lib/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=0.9313543599257884, f1=0.919422449930135,
best_f1=0.919422449930135
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.026051972061395645
step: 10, loss: 0.024892106652259827
step: 20, loss: 0.1338319033384323
step: 30, loss: 0.13658416271209717
step: 40, loss: 0.024407079443335533
step: 50, loss: 0.05785354599356651
step: 60, loss: 0.017411643639206886
step: 70, loss: 0.010372638702392578
step: 80, loss: 0.02464539371430874
step: 90, loss: 0.03965340182185173
step: 100, loss: 0.0231808852404356
step: 110, loss: 0.10646820068359375
step: 120, loss: 0.02037741243839264
step: 130, loss: 0.013230687938630581
step: 140, loss: 0.0805291160941124
step: 150, loss: 0.0098115224391222
step: 160, loss: 0.0020509944297373295
step: 170, loss: 0.01859535090625286
step: 180, loss: 0.007011936977505684
step: 190, loss: 0.013156427070498466
step: 200, loss: 0.06100895628333092
step: 210, loss: 0.0885562151670456
step: 220, loss: 0.050496116280555725
step: 230, loss: 0.021622976288199425
step: 240, loss: 0.09812527149915695
step: 250, loss: 0.01905796304345131
step: 260, loss: 0.001438023871742189
step: 270, loss: 0.0071294354274868965
step: 280, loss: 0.015124017372727394
step: 290, loss: 0.1135200560092926
step: 300, loss: 0.030092982575297356
step: 310, loss: 0.07663147151470184
step: 320, loss: 0.09524040669202805
step: 330, loss: 0.003154444508254528
step: 340, loss: 0.011040850542485714
step: 350, loss: 0.009687142446637154
step: 360, loss: 0.01497221365571022
step: 370, loss: 0.01424307469278574
step: 380, loss: 0.003588481340557337
step: 390, loss: 0.04209626093506813
step: 400, loss: 0.021007388830184937
step: 410, loss: 0.01493427436798811
step: 420, loss: 0.08337662369012833
step: 430, loss: 0.04810218885540962
step: 440, loss: 0.005911197979003191
step: 450, loss: 0.031057005748152733
```

```
step: 460, loss: 0.07795373350381851
step: 470, loss: 0.004944221582263708
step: 480, loss: 0.005238296929746866
step: 490, loss: 0.0708877220749855
step: 500, loss: 0.06139276549220085
step: 510, loss: 0.007854390889406204
step: 520, loss: 0.006291849538683891
step: 530, loss: 0.013963536359369755
/usr/local/lib/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=0.9318600368324126, f1=0.9232914923291492,
best_f1=0.9232914923291492
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04228924214839935
step: 10, loss: 0.01714225299656391
step: 20, loss: 0.010701206512749195
step: 30, loss: 0.020261092111468315
step: 40, loss: 0.010300333611667156
step: 50, loss: 0.011359135620296001
step: 60, loss: 0.0004971657763235271
step: 70, loss: 0.002459977986291051
step: 80, loss: 0.0020789573900401592
step: 90, loss: 0.00032134252251125877
step: 100, loss: 0.011315747164189816
step: 110, loss: 0.0002974280796479434
step: 120, loss: 0.0005384381511248648
step: 130, loss: 0.03398151323199272
step: 140, loss: 0.0007209215546026826
step: 150, loss: 0.0012178717879578471
step: 160, loss: 0.0011600800789892673
step: 170, loss: 0.0007407237426377833
step: 180, loss: 0.04245935007929802
step: 190, loss: 0.010874895378947258
step: 200, loss: 0.002026498317718506
step: 210, loss: 0.02384316734969616
step: 220, loss: 0.017083579674363136
step: 230, loss: 0.20870080590248108
step: 240, loss: 0.018500395119190216
step: 250, loss: 0.00541676813736558
step: 260, loss: 0.05798165872693062
step: 270, loss: 0.0867883637547493
```

```
step: 280, loss: 0.004313153214752674
step: 290, loss: 0.12298581749200821
step: 300, loss: 0.007028837222605944
step: 310, loss: 0.02851354330778122
step: 320, loss: 0.0064874934032559395
step: 330, loss: 0.012971206568181515
step: 340, loss: 0.02159424126148224
step: 350, loss: 0.07278607785701752
step: 360, loss: 0.03731277585029602
step: 370, loss: 0.04475303366780281
step: 380, loss: 0.0034673353657126427
step: 390, loss: 0.12297667562961578
step: 400, loss: 0.008378883823752403
step: 410, loss: 0.02701796218752861
step: 420, loss: 0.010967250913381577
step: 430, loss: 0.009081032127141953
step: 440, loss: 0.04336994141340256
step: 450, loss: 0.004344182554632425
step: 460, loss: 0.001766903791576624
step: 470, loss: 0.004845276474952698
step: 480, loss: 0.0019772795494645834
step: 490, loss: 0.040962815284729004
step: 500, loss: 0.005795951467007399
step: 510, loss: 0.006213682238012552
step: 520, loss: 0.07136201113462448
step: 530, loss: 0.0006530591635964811
/usr/local/lib/python3.7/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=0.9242424242424242, f1=0.9107398568019094,
best_f1=0.9232914923291492
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0038520197849720716
step: 10, loss: 0.0030361393000930548
step: 20, loss: 0.0017811321886256337
step: 30, loss: 0.005370995961129665
step: 40, loss: 0.013597636483609676
step: 50, loss: 0.03903978690505028
step: 60, loss: 0.006802701856940985
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 70, loss: 0.08450116962194443
step: 80, loss: 0.049637194722890854
step: 90, loss: 0.006353430449962616
step: 100, loss: 0.011135192587971687
step: 110, loss: 0.029466383159160614
step: 120, loss: 0.005531201604753733
```

```
step: 130, loss: 0.0586879663169384
step: 140, loss: 0.0007324175094254315
step: 150, loss: 0.0007244684966281056
step: 160, loss: 0.05117348954081535
step: 170, loss: 0.0018056068802252412
step: 180, loss: 0.1811576783657074
step: 190, loss: 0.0023669335059821606
step: 200, loss: 0.0007955709006637335
step: 210, loss: 0.0030062971636652946
step: 220, loss: 0.00031302921706810594
step: 230, loss: 0.0009526566718704998
step: 240, loss: 0.0037391120567917824
step: 250, loss: 0.0011030228342860937
step: 260, loss: 0.0030390098690986633
step: 270, loss: 0.009401874616742134
step: 280, loss: 0.011096087284386158
step: 290, loss: 0.0007599960663355887
step: 300, loss: 0.058122411370277405
step: 310, loss: 0.0043784016743302345
step: 320, loss: 0.0024319873191416264
step: 330, loss: 0.00217342097312212
step: 340, loss: 0.0007886621751822531
step: 350, loss: 0.09368070960044861
step: 360, loss: 0.0023405307438224554
step: 370, loss: 0.014073533937335014
step: 380, loss: 0.028944306075572968
step: 390, loss: 0.000836795661598444
step: 400, loss: 0.11209461838006973
step: 410, loss: 0.0003999779000878334
step: 420, loss: 0.0033239710610359907
step: 430, loss: 0.002776741050183773
step: 440, loss: 0.01446880679577589
step: 450, loss: 0.006991564761847258
step: 460, loss: 0.0024749746080487967
step: 470, loss: 0.048299141228199005
step: 480, loss: 0.015316102653741837
step: 490, loss: 0.006072274409234524
step: 500, loss: 0.001506010303273797
step: 510, loss: 0.002216164954006672
step: 520, loss: 0.0010305900359526277
step: 530, loss: 0.00725582055747509
/usr/local/lib/python3.7/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=0.9246417013407305, f1=0.9147647880763857,
best_f1=0.9232914923291492
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021415906958281994
step: 10, loss: 0.0022485163062810898
step: 20, loss: 0.013918040320277214
step: 30, loss: 0.004168884363025427
step: 40, loss: 0.0002319843479199335
step: 50, loss: 0.001748319249600172
step: 60, loss: 0.00012725725537166
step: 70, loss: 0.021874576807022095
step: 80, loss: 0.0014703389024361968
step: 90, loss: 0.1497783660888672
step: 100, loss: 0.006126490421593189
step: 110, loss: 0.004023509565740824
step: 120, loss: 0.000601961393840611
step: 130, loss: 0.0002461168041918427
step: 140, loss: 0.0007662408170290291
step: 150, loss: 0.0011216767597943544
step: 160, loss: 0.057543691247701645
step: 170, loss: 0.0003114765859209001
step: 180, loss: 0.0002836525090970099
step: 190, loss: 0.0005860573728568852
step: 200, loss: 0.00036360096419230103
step: 210, loss: 0.00035293662222102284
step: 220, loss: 0.0002852336911018938
step: 230, loss: 0.0003701960376929492
step: 240, loss: 0.00022390903905034065
step: 250, loss: 0.0001958796347025782
step: 260, loss: 0.0002984458697028458
step: 270, loss: 8.323867223225534e-05
step: 280, loss: 0.028935126960277557
step: 290, loss: 0.00019469826656859368
step: 300, loss: 0.0010682175634428859
step: 310, loss: 0.0008504827274009585
step: 320, loss: 0.1465967893600464
step: 330, loss: 0.002405308187007904
step: 340, loss: 0.024096792563796043
step: 350, loss: 0.04545450210571289
step: 360, loss: 0.020246131345629692
step: 370, loss: 0.0009066346101462841
step: 380, loss: 0.0009796347003430128
step: 390, loss: 0.015262903645634651
step: 400, loss: 0.00017835857579484582
step: 410, loss: 0.00019632221665233374
step: 420, loss: 0.0009694250184111297
step: 430, loss: 0.0006267401622608304
step: 440, loss: 0.002506654942408204
step: 450, loss: 0.0011585184838622808
step: 460, loss: 0.07415053248405457
```

```
step: 470, loss: 0.032961536198854446
step: 480, loss: 0.005319529213011265
step: 490, loss: 0.00016304090968333185
step: 500, loss: 0.004842095077037811
step: 510, loss: 0.0008731495472602546
step: 520, loss: 0.00033782096579670906
step: 530, loss: 0.0025970523711293936
/usr/local/lib/python3.7/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=0.925627664613927, f1=0.9103053435114504,
best_f1=0.9232914923291492
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011342735029757023
step: 10, loss: 0.004477169830352068
step: 20, loss: 0.009373062290251255
step: 30, loss: 0.0011411524610593915
step: 40, loss: 0.0007908347761258483
step: 50, loss: 0.0004401701153255999
step: 60, loss: 0.0012233416782692075
step: 70, loss: 0.0010823990451171994
step: 80, loss: 0.00026360811898484826
step: 90, loss: 0.0003116318548563868
step: 100, loss: 0.004440205171704292
step: 110, loss: 0.006735872011631727
step: 120, loss: 0.001694549573585391
step: 130, loss: 0.002068039495497942
step: 140, loss: 0.0004639727994799614
step: 150, loss: 0.00014028909208718687
step: 160, loss: 0.0017123478464782238
step: 170, loss: 0.000991349690593779
step: 180, loss: 0.0011703006457537413
step: 190, loss: 0.00040844734758138657
step: 200, loss: 0.0012453782837837934
step: 210, loss: 0.010919437743723392
step: 220, loss: 7.954129978315905e-05
step: 230, loss: 0.00019322906155139208
step: 240, loss: 0.023486025631427765
step: 250, loss: 0.00038624630542472005
step: 260, loss: 0.0006173423607833683
step: 270, loss: 0.0004745852202177048
step: 280, loss: 0.01202467456459999
step: 290, loss: 0.0036237789317965508
step: 300, loss: 0.00011852538591483608
step: 310, loss: 0.003934777341783047
step: 320, loss: 0.055356789380311966
```

```
step: 330, loss: 0.002874396275728941
step: 340, loss: 0.047778405249118805
step: 350, loss: 0.002654485171660781
step: 360, loss: 0.0008540608105249703
step: 370, loss: 0.00018427478789817542
step: 380, loss: 0.00013781260349787772
step: 390, loss: 5.930526458541863e-05
step: 400, loss: 5.296360905049369e-05
step: 410, loss: 0.0007520415238104761
step: 420, loss: 0.0001012614302453585
step: 430, loss: 0.00023737846640869975
step: 440, loss: 0.00691950274631381
step: 450, loss: 0.00027659276383928955
step: 460, loss: 0.00019313265511300415
step: 470, loss: 0.018952246755361557
step: 480, loss: 0.00020646305347327143
step: 490, loss: 6.761474651284516e-05
step: 500, loss: 0.0001912508305395022
step: 510, loss: 0.0003826030297204852
step: 520, loss: 0.00022223088308237493
step: 530, loss: 6.454369577113539e-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=0.9255121042830541, f1=0.9172932330827067,
best_f1=0.9232914923291492
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021394956274888941
step: 10, loss: 0.0014179437421262264
step: 20, loss: 0.0021272923331707716
step: 30, loss: 0.0003733767371159047
step: 40, loss: 0.00011376183829270303
step: 50, loss: 0.00028023013146594167
step: 60, loss: 0.00010036843013949692
step: 70, loss: 0.05772740766406059
step: 80, loss: 5.465824870043434e-05
step: 90, loss: 7.775864651193842e-05
step: 100, loss: 0.0005969173507764935
step: 110, loss: 0.00032463009119965136
step: 120, loss: 0.017104851081967354
step: 130, loss: 0.0003751225012820214
step: 140, loss: 4.262324728188105e-05
step: 150, loss: 0.0005309860571287572
step: 160, loss: 0.02687792479991913
step: 170, loss: 0.0027168835513293743
step: 180, loss: 0.0002341526560485363
```

```
step: 190, loss: 0.000799870234914124
step: 200, loss: 0.0007192856282927096
step: 210, loss: 0.03849292919039726
step: 220, loss: 0.03536376729607582
step: 230, loss: 0.14129436016082764
step: 240, loss: 0.0009224979439750314
step: 250, loss: 0.0068215844221413136
step: 260, loss: 0.0007425493677146733
step: 270, loss: 0.00018407120660413057
step: 280, loss: 0.0013670255430042744
step: 290, loss: 0.0007722426089458168
step: 300, loss: 0.0002771865401882678
step: 310, loss: 0.01681353896856308
step: 320, loss: 0.00026589990011416376
step: 330, loss: 0.0012495265109464526
step: 340, loss: 0.00013673347712028772
step: 350, loss: 0.003668698016554117
step: 360, loss: 0.00028901148471049964
step: 370, loss: 0.00019489358237478882
step: 380, loss: 0.039373401552438736
step: 390, loss: 0.0009703740943223238
step: 400, loss: 0.03971562907099724
step: 410, loss: 0.0005254787974990904
step: 420, loss: 0.00020045287965331227
step: 430, loss: 0.0007378109148703516
step: 440, loss: 0.0002314118028152734
step: 450, loss: 0.00035862118238583207
step: 460, loss: 0.003184955334290862
step: 470, loss: 0.000263408845057711
step: 480, loss: 8.659984450787306e-05
step: 490, loss: 0.0002519116969779134
step: 500, loss: 0.0003976819571107626
step: 510, loss: 9.887929627439007e-05
step: 520, loss: 0.00018855567032005638
step: 530, loss: 0.00010712318180594593
/usr/local/lib/python3.7/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=0.9267605633802817, f1=0.906679298910469,
best_f1=0.9232914923291492
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001460541971027851
step: 10, loss: 0.00013426286750473082
step: 20, loss: 0.00012894286192022264
step: 30, loss: 0.0009965661447495222
step: 40, loss: 0.00010592555190669373
```

```
step: 50, loss: 0.00019067636458203197
step: 60, loss: 0.00012802373385056853
step: 70, loss: 0.011199578642845154
step: 80, loss: 7.264983287313953e-05
step: 90, loss: 0.0017019223887473345
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 100, loss: 0.009916547685861588
step: 110, loss: 8.882765541784465e-05
step: 120, loss: 0.0002154761168640107
step: 130, loss: 0.0673057809472084
step: 140, loss: 0.00014848847058601677
step: 150, loss: 0.00045692423009313643
step: 160, loss: 8.622970199212432e-05
step: 170, loss: 0.0033916416577994823
step: 180, loss: 5.304874503053725e-05
step: 190, loss: 0.021351447328925133
step: 200, loss: 0.00040965521475300193
step: 210, loss: 4.802467447007075e-05
step: 220, loss: 0.003778276266530156
step: 230, loss: 4.7167075535980985e-05
step: 240, loss: 0.00027560352464206517
step: 250, loss: 0.0355515219271183
step: 260, loss: 0.00011821813677670434
step: 270, loss: 0.001846848987042904
step: 280, loss: 0.00010903174552368
step: 290, loss: 7.642417767783627e-05
step: 300, loss: 0.001629821490496397
step: 310, loss: 5.385169424698688e-05
step: 320, loss: 0.01058455090969801
step: 330, loss: 0.004606177564710379
step: 340, loss: 6.762074190191925e-05
step: 350, loss: 7.000335608609021e-05
step: 360, loss: 0.011830538511276245
step: 370, loss: 0.00013906268577557057
step: 380, loss: 5.21008187206462e-05
step: 390, loss: 0.00012182525097159669
step: 400, loss: 0.007337084040045738
step: 410, loss: 0.001854276517406106
step: 420, loss: 0.00018657359760254622
step: 430, loss: 5.5089392844820395e-05
step: 440, loss: 8.574616367695853e-05
step: 450, loss: 8.351082942681387e-05
step: 460, loss: 0.0003390669880900532
step: 470, loss: 0.00037686043651774526
step: 480, loss: 0.002443664940074086
step: 490, loss: 0.028494572266936302
step: 500, loss: 0.0010507473489269614
step: 510, loss: 6.45047621219419e-05
```

```
step: 520, loss: 0.0031001416500657797
step: 530, loss: 0.0015770914033055305
/usr/local/lib/python3.7/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=0.9279112754158966, f1=0.9127579737335835,
best_f1=0.9232914923291492
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007594256312586367
step: 10, loss: 5.594481626758352e-05
step: 20, loss: 0.0003298044903203845
step: 30, loss: 8.24410526547581e-05
step: 40, loss: 0.00031115987803786993
step: 50, loss: 0.00017396766634192318
step: 60, loss: 0.00017693062545731664
step: 70, loss: 0.00029144983272999525
step: 80, loss: 0.0010261830175295472
step: 90, loss: 0.00011601961159612983
step: 100, loss: 0.0007496471516788006
step: 110, loss: 0.00026040751254186034
step: 120, loss: 8.761997014516965e-05
step: 130, loss: 9.68955791904591e-05
step: 140, loss: 0.01865481771528721
step: 150, loss: 6.317066436167806e-05
step: 160, loss: 0.0018500767182558775
step: 170, loss: 0.03799045830965042
step: 180, loss: 0.002839325927197933
step: 190, loss: 0.00028303329600021243
step: 200, loss: 0.0014945995062589645
step: 210, loss: 0.000350289570633322
step: 220, loss: 0.0001230499765370041
step: 230, loss: 0.0001602064148755744
step: 240, loss: 0.00020176857651676983
step: 250, loss: 0.0008948769536800683
step: 260, loss: 0.00026471898308955133
step: 270, loss: 4.292879020795226e-05
step: 280, loss: 7.793093391228467e-05
step: 290, loss: 3.825745443464257e-05
step: 300, loss: 0.0001120470988098532
step: 310, loss: 0.00012264570978004485
step: 320, loss: 6.215695611899719e-05
step: 330, loss: 7.819577876944095e-05
step: 340, loss: 5.24630413565319e-05
step: 350, loss: 0.0005736741004511714
step: 360, loss: 0.00017119795666076243
step: 370, loss: 0.00017433782340958714
```

```
step: 380, loss: 0.0009221512591466308
step: 390, loss: 8.376890036743134e-05
step: 400, loss: 0.0008163820602931082
step: 410, loss: 0.0002611708187032491
step: 420, loss: 0.0005509315524250269
step: 430, loss: 3.6721703509101644e-05
step: 440, loss: 4.366154462331906e-05
step: 450, loss: 0.0002148423227481544
step: 460, loss: 0.00018241204088553786
step: 470, loss: 0.0002651884569786489
step: 480, loss: 0.00043207156704738736
step: 490, loss: 0.044008925557136536
step: 500, loss: 0.0002908113820012659
step: 510, loss: 0.0002445194695610553
step: 520, loss: 0.00037449688534252346
step: 530, loss: 0.007248105015605688
/usr/local/lib/python3.7/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=0.9315838800374883, f1=0.9153664302600472,
best_f1=0.9232914923291492
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.434260310721584e-05
step: 10, loss: 0.00025069425464607775
step: 20, loss: 2.810275873343926e-05
step: 30, loss: 6.477193528553471e-05
step: 40, loss: 0.00015143811469897628
step: 50, loss: 0.011681224219501019
step: 60, loss: 0.002814943203702569
step: 70, loss: 0.00012333111953921616
step: 80, loss: 0.09523118287324905
step: 90, loss: 0.00024033902445808053
step: 100, loss: 2.5893968995660543e-05
step: 110, loss: 0.0001313328684773296
step: 120, loss: 0.0002777176268864423
step: 130, loss: 9.54400165937841e-05
step: 140, loss: 0.00016622668772470206
step: 150, loss: 4.801595423487015e-05
step: 160, loss: 0.0002836313215084374
step: 170, loss: 0.0007776482962071896
step: 180, loss: 4.365807035355829e-05
step: 190, loss: 0.0003566728555597365
step: 200, loss: 0.00032468963763676584
step: 210, loss: 0.00032037400524131954
step: 220, loss: 7.614616333739832e-05
step: 230, loss: 6.364403816405684e-05
```

```
step: 240, loss: 7.574634219054133e-05
step: 250, loss: 6.476158159784973e-05
step: 260, loss: 0.0016021106857806444
step: 270, loss: 0.0005070175975561142
step: 280, loss: 0.015999101102352142
step: 290, loss: 0.009509407915174961
step: 300, loss: 0.01113712228834629
step: 310, loss: 0.000666289939545095
step: 320, loss: 9.38258963287808e-05
step: 330, loss: 0.0028025307692587376
step: 340, loss: 4.71280436613597e-05
step: 350, loss: 5.182723543839529e-05
step: 360, loss: 0.0010093317832797766
step: 370, loss: 3.6711608117911965e-05
step: 380, loss: 0.00014928918972145766
step: 390, loss: 0.0018871997017413378
step: 400, loss: 0.00014182153972797096
step: 410, loss: 8.418416109634563e-05
step: 420, loss: 0.0005795530159957707
step: 430, loss: 0.0002446794824209064
step: 440, loss: 0.00015301084204111248
step: 450, loss: 0.0002798394998535514
step: 460, loss: 0.00030802187393419445
step: 470, loss: 0.0006326739094220102
step: 480, loss: 4.250747952028178e-05
step: 490, loss: 0.0009017562260851264
step: 500, loss: 0.00010180578829022124
step: 510, loss: 8.492046617902815e-05
step: 520, loss: 3.111662226729095e-05
step: 530, loss: 2.625151137181092e-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=0.9298245614035088, f1=0.9172093023255814,
best_f1=0.9232914923291492
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004163255216553807
step: 10, loss: 3.996724262833595e-05
step: 20, loss: 2.736165333772078e-05
step: 30, loss: 4.476876347325742e-05
step: 40, loss: 9.624616359360516e-05
step: 50, loss: 3.543734783306718e-05
step: 60, loss: 0.0003345143923070282
step: 70, loss: 0.00023192583466880023
step: 80, loss: 0.0010529137216508389
step: 90, loss: 0.00019102089572697878
```

```
step: 100, loss: 0.001416516024619341
step: 110, loss: 4.44102224719245e-05
step: 120, loss: 0.00013543629029300064
step: 130, loss: 6.252004823181778e-05
step: 140, loss: 4.1973184124799445e-05
step: 150, loss: 0.00028836025740019977
step: 160, loss: 0.0003217019257135689
step: 170, loss: 0.0006973369471415877
step: 180, loss: 5.4065341828390956e-05
step: 190, loss: 2.0462692191358656e-05
step: 200, loss: 4.391880429466255e-05
step: 210, loss: 6.19174461462535e-05
step: 220, loss: 7.064594683470204e-05
step: 230, loss: 5.220533785177395e-05
step: 240, loss: 6.002940426697023e-05
step: 250, loss: 0.00012225071259308606
step: 260, loss: 0.00012103764311177656
step: 270, loss: 4.608596645994112e-05
step: 280, loss: 5.496077938005328e-05
step: 290, loss: 0.0002535535313654691
step: 300, loss: 0.0005577486590482295
step: 310, loss: 0.0011265368666499853
step: 320, loss: 4.360774255474098e-05
step: 330, loss: 4.0152346628019586e-05
step: 340, loss: 9.054612019099295e-05
step: 350, loss: 0.00238824900612235
step: 360, loss: 4.552337122731842e-05
step: 370, loss: 3.0192297344910912e-05
step: 380, loss: 0.00011855359480250627
step: 390, loss: 3.371997445356101e-05
step: 400, loss: 5.832393799209967e-05
step: 410, loss: 7.606685539940372e-05
step: 420, loss: 0.001079097273759544
step: 430, loss: 0.02912786416709423
step: 440, loss: 4.66218261863105e-05
step: 450, loss: 0.015517045743763447
step: 460, loss: 7.123070827219635e-05
step: 470, loss: 4.598001396516338e-05
step: 480, loss: 7.813285628799349e-05
step: 490, loss: 0.00020390027202665806
step: 500, loss: 0.021643688902258873
step: 510, loss: 0.0015898386482149363
step: 520, loss: 0.0003734836936928332
step: 530, loss: 0.00019736445392481983
/usr/local/lib/python3.7/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=0.9290681502086231, f1=0.9161592505854801,
```

```
best_f1=0.9232914923291492
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.775488534709439e-05
step: 10, loss: 3.1010764359962195e-05
step: 20, loss: 4.2390209273435175e-05
step: 30, loss: 6.195328023750335e-05
step: 40, loss: 2.5331337383249775e-05
step: 50, loss: 2.9190454370109364e-05
step: 60, loss: 6.105343345552683e-05
step: 70, loss: 2.7033358492190018e-05
step: 80, loss: 4.5350891014095396e-05
step: 90, loss: 5.431679164757952e-05
step: 100, loss: 2.2593199901166372e-05
step: 110, loss: 4.143719343119301e-05
step: 120, loss: 8.309905388159677e-05
step: 130, loss: 7.205833389889449e-05
step: 140, loss: 3.3093678212026134e-05
step: 150, loss: 0.00011208268551854417
step: 160, loss: 2.7540381779544987e-05
step: 170, loss: 0.00012649784912355244
step: 180, loss: 3.087021104875021e-05
step: 190, loss: 4.892547440249473e-05
step: 200, loss: 0.003086183685809374
step: 210, loss: 0.0008845381671562791
step: 220, loss: 2.3595588572788984e-05
step: 230, loss: 2.8083944926038384e-05
step: 240, loss: 3.592521898099221e-05
step: 250, loss: 0.0029660016298294067
step: 260, loss: 0.00019724344019778073
step: 270, loss: 3.1343402952188626e-05
step: 280, loss: 2.744659832387697e-05
step: 290, loss: 5.5853419326012954e-05
step: 300, loss: 9.630585554987192e-05
step: 310, loss: 0.0003112433187197894
step: 320, loss: 3.654275860753842e-05
step: 330, loss: 3.1473959097638726e-05
step: 340, loss: 2.985730316140689e-05
step: 350, loss: 0.0009768034797161818
step: 360, loss: 0.0008382105734199286
step: 370, loss: 8.296383748529479e-05
step: 380, loss: 4.12957051594276e-05
step: 390, loss: 3.9798895159037784e-05
step: 400, loss: 2.416515962977428e-05
step: 410, loss: 3.443758760113269e-05
step: 420, loss: 4.313061435823329e-05
step: 430, loss: 0.0004977190401405096
```

```
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 440, loss: 0.0001585962891113013
step: 450, loss: 3.6696237657452e-05
step: 460, loss: 2.3274902559933253e-05
step: 470, loss: 0.000673813046887517
step: 480, loss: 0.00021977971482556313
step: 490, loss: 0.00013761635636910796
step: 500, loss: 7.656441448489204e-05
step: 510, loss: 9.641791984904557e-05
step: 520, loss: 2.5264203941333108e-05
step: 530, loss: 3.967126031056978e-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=0.9304467987102718, f1=0.9197761194029851,
best_f1=0.9232914923291492
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.158363273949362e-05
step: 10, loss: 2.9142176572349854e-05
step: 20, loss: 3.9318572817137465e-05
step: 30, loss: 3.3344105759169906e-05
step: 40, loss: 2.4686898541403934e-05
step: 50, loss: 2.416918687231373e-05
step: 60, loss: 0.00012355744547676295
step: 70, loss: 2.17774577322416e-05
step: 80, loss: 3.21816150972154e-05
step: 90, loss: 1.699082349659875e-05
step: 100, loss: 4.6139823098201305e-05
step: 110, loss: 2.114435483235866e-05
step: 120, loss: 0.00016114788013510406
step: 130, loss: 1.9929975678678602e-05
step: 140, loss: 0.0008125349413603544
step: 150, loss: 0.00011508030002005398
step: 160, loss: 0.00043832394294440746
step: 170, loss: 9.366836457047611e-05
step: 180, loss: 5.059429531684145e-05
step: 190, loss: 4.9844129534903914e-05
step: 200, loss: 3.440884393057786e-05
step: 210, loss: 0.00033324683317914605
step: 220, loss: 3.345350341987796e-05
step: 230, loss: 0.0003623106167651713
step: 240, loss: 6.23802188783884e-05
step: 250, loss: 3.471425225143321e-05
step: 260, loss: 1.7743310309015214e-05
step: 270, loss: 0.0012422794243320823
step: 280, loss: 2.3699592929915525e-05
```

```
step: 290, loss: 8.42489316710271e-05
step: 300, loss: 0.0005328389233909547
step: 310, loss: 2.0496092474786565e-05
step: 320, loss: 2.6404251912026666e-05
step: 330, loss: 3.159579500788823e-05
step: 340, loss: 0.0002469006576575339
step: 350, loss: 0.00011685440404107794
step: 360, loss: 0.00019287194299977273
step: 370, loss: 2.9204802558524534e-05
step: 380, loss: 4.6442317398032174e-05
step: 390, loss: 2.7789294108515605e-05
step: 400, loss: 3.544108403730206e-05
step: 410, loss: 4.722396261058748e-05
step: 420, loss: 5.3528769058175385e-05
step: 430, loss: 2.3032982426229864e-05
step: 440, loss: 2.5677494704723358e-05
step: 450, loss: 2.8321521313046105e-05
step: 460, loss: 0.00014865331468172371
step: 470, loss: 1.8898055714089423e-05
step: 480, loss: 3.925767305190675e-05
step: 490, loss: 5.06782962474972e-05
step: 500, loss: 9.90421322057955e-05
step: 510, loss: 2.688031418074388e-05
step: 520, loss: 0.0008799495990388095
step: 530, loss: 2.2749900381313637e-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 14: dev_f1=0.9324200913242009, f1=0.9194661757938334,
best_f1=0.9194661757938334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.767249290831387e-05
step: 10, loss: 0.04449464753270149
step: 20, loss: 2.460126779624261e-05
step: 30, loss: 0.00254362472333014
step: 40, loss: 1.5642242942703888e-05
step: 50, loss: 4.101316153537482e-05
step: 60, loss: 0.00020741105254273862
step: 70, loss: 8.968254405772313e-05
step: 80, loss: 5.7825112889986485e-05
step: 90, loss: 4.588699812302366e-05
step: 100, loss: 1.6957295883912593e-05
```

```
step: 110, loss: 0.00026880792574957013
step: 120, loss: 0.0007203003624454141
step: 130, loss: 1.631647683097981e-05
step: 140, loss: 1.6458092432003468e-05
step: 150, loss: 0.00022322387667372823
step: 160, loss: 3.339412432978861e-05
step: 170, loss: 0.00022086972603574395
step: 180, loss: 4.3735566578106955e-05
step: 190, loss: 0.0006142742349766195
step: 200, loss: 2.6292434995411895e-05
step: 210, loss: 6.756801303708926e-05
step: 220, loss: 1.740050174703356e-05
step: 230, loss: 1.976599742192775e-05
step: 240, loss: 2.9861121220164932e-05
step: 250, loss: 4.036620885017328e-05
step: 260, loss: 2.9484321203199215e-05
step: 270, loss: 0.0001657532120589167
step: 280, loss: 9.042176679940894e-05
step: 290, loss: 0.002934193005785346
step: 300, loss: 0.0008657114813104272
step: 310, loss: 5.895541107747704e-05
step: 320, loss: 2.7201156626688316e-05
step: 330, loss: 0.00031953194411471486
step: 340, loss: 4.9536469305166975e-05
step: 350, loss: 0.00015937426360324025
step: 360, loss: 2.211666651419364e-05
step: 370, loss: 1.5616244127159007e-05
step: 380, loss: 3.175896927132271e-05
step: 390, loss: 2.483146636222955e-05
step: 400, loss: 2.2216723664314486e-05
step: 410, loss: 3.601061325753108e-05
step: 420, loss: 6.856917752884328e-05
step: 430, loss: 1.9870361938956194e-05
step: 440, loss: 0.0001705258182482794
step: 450, loss: 0.00016388250514864922
step: 460, loss: 3.205378379789181e-05
step: 470, loss: 4.2657491576392204e-05
step: 480, loss: 2.0648854842875153e-05
step: 490, loss: 3.72265130863525e-05
step: 500, loss: 3.0147904908517376e-05
step: 510, loss: 0.0013645951403304935
step: 520, loss: 2.4921369913499802e-05
step: 530, loss: 2.47353073064005e-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=0.9314312011044639, f1=0.9212121212121211,
best_f1=0.9194661757938334
```

```
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.")
5742it [00:17, 321.18it/s]
load_f1 = 0.9297752808988764
real_f1 = 0.9279700654817586
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, 369.74it/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   : 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: 0.8301610946655273
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.06135062128305435
step: 20, loss: 0.38504961133003235
step: 30, loss: 0.3765375018119812
step: 40, loss: 0.5139895677566528
step: 50, loss: 0.3013783097267151
step: 60, loss: 0.37102609872817993
step: 70, loss: 0.2726202607154846
step: 80, loss: 0.2834383547306061
step: 90, loss: 0.4517812430858612
step: 100, loss: 0.19246940314769745
step: 110, loss: 0.28356051445007324
step: 120, loss: 0.24802866578102112
step: 130, loss: 0.2220546454191208
step: 140, loss: 0.2667624354362488
step: 150, loss: 0.2868172228336334
step: 160, loss: 0.2505057454109192
step: 170, loss: 0.19373512268066406
step: 180, loss: 0.23985961079597473
step: 190, loss: 0.21463455259799957
step: 200, loss: 0.19190925359725952
step: 210, loss: 0.4954773783683777
/usr/local/lib/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=0.5136612021857923, f1=0.5108055009823184,
best_f1=0.5108055009823184
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08287351578474045
step: 10, loss: 0.06906978040933609
step: 20, loss: 0.22037538886070251
step: 30, loss: 0.20537279546260834
step: 40, loss: 0.019227124750614166
step: 50, loss: 0.20474091172218323
step: 60, loss: 0.0727265328168869
step: 70, loss: 0.23091720044612885
step: 80, loss: 0.2868408262729645
step: 90, loss: 0.13837449252605438
step: 100, loss: 0.11732129007577896
step: 110, loss: 0.09718241542577744
step: 120, loss: 0.2001543641090393
step: 130, loss: 0.273051917552948
step: 140, loss: 0.28984200954437256
step: 150, loss: 0.2266402691602707
step: 160, loss: 0.16136106848716736
step: 170, loss: 0.1815851479768753
step: 180, loss: 0.2600695490837097
step: 190, loss: 0.2282407581806183
step: 200, loss: 0.12290878593921661
step: 210, loss: 0.2326129823923111
/usr/local/lib/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=0.5724907063197026, f1=0.6015625, best_f1=0.6015625
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2418418973684311
step: 10, loss: 0.15664251148700714
step: 20, loss: 0.23863404989242554
step: 30, loss: 0.10001676529645592
step: 40, loss: 0.2128378003835678
step: 50, loss: 0.13941410183906555
step: 60, loss: 0.17727147042751312
step: 70, loss: 0.06735670566558838
step: 80, loss: 0.15523983538150787
step: 90, loss: 0.11348927766084671
step: 100, loss: 0.03765793889760971
step: 110, loss: 0.15590420365333557
step: 120, loss: 0.11211372911930084
```

```
step: 130, loss: 0.1587090790271759
step: 140, loss: 0.34574776887893677
step: 150, loss: 0.13258476555347443
step: 160, loss: 0.13061918318271637
step: 170, loss: 0.12100962549448013
step: 180, loss: 0.16227683424949646
step: 190, loss: 0.31345534324645996
step: 200, loss: 0.13771851360797882
step: 210, loss: 0.3013119697570801
/usr/local/lib/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=0.5925925925925926, f1=0.5987525987525987,
best_f1=0.5987525987525987
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21697315573692322
step: 10, loss: 0.06678295880556107
step: 20, loss: 0.043819278478622437
step: 30, loss: 0.09952924400568008
step: 40, loss: 0.07852805405855179
step: 50, loss: 0.08730711042881012
step: 60, loss: 0.048832159489393234
step: 70, loss: 0.1463356912136078
step: 80, loss: 0.17504705488681793
step: 90, loss: 0.06362570077180862
step: 100, loss: 0.1849449723958969
step: 110, loss: 0.16725178062915802
step: 120, loss: 0.12382999807596207
step: 130, loss: 0.11439727991819382
step: 140, loss: 0.09061781316995621
step: 150, loss: 0.16973723471164703
step: 160, loss: 0.07496815919876099
step: 170, loss: 0.08571263402700424
step: 180, loss: 0.2686239778995514
step: 190, loss: 0.07367311418056488
step: 200, loss: 0.16896456480026245
step: 210, loss: 0.08468256890773773
/usr/local/lib/python3.7/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=0.5604838709677419, f1=0.5546218487394957,
best_f1=0.5987525987525987
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05196104571223259
step: 10, loss: 0.08352115005254745
step: 20, loss: 0.01633637771010399
step: 30, loss: 0.030692722648382187
step: 40, loss: 0.05649831146001816
step: 50, loss: 0.11433664709329605
step: 60, loss: 0.08127988874912262
step: 70, loss: 0.03668248653411865
step: 80, loss: 0.018055645748972893
step: 90, loss: 0.11431022733449936
step: 100, loss: 0.07428716123104095
step: 110, loss: 0.007268047425895929
step: 120, loss: 0.08234816044569016
step: 130, loss: 0.1004752442240715
step: 140, loss: 0.09428998827934265
step: 150, loss: 0.03126335144042969
step: 160, loss: 0.05074786767363548
step: 170, loss: 0.03909610956907272
step: 180, loss: 0.09385048598051071
step: 190, loss: 0.10744783282279968
step: 200, loss: 0.18576303124427795
step: 210, loss: 0.132281094789505
/usr/local/lib/python3.7/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=0.5649717514124294, f1=0.5508982035928144,
best_f1=0.5987525987525987
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.059879373759031296
step: 10, loss: 0.07378469407558441
step: 20, loss: 0.041082095354795456
step: 30, loss: 0.014033801853656769
step: 40, loss: 0.061376046389341354
step: 50, loss: 0.02380654774606228
step: 60, loss: 0.07815809547901154
step: 70, loss: 0.027229705825448036
step: 80, loss: 0.13021402060985565
step: 90, loss: 0.1099391058087349
step: 100, loss: 0.0467662550508976
step: 110, loss: 0.009722575545310974
step: 120, loss: 0.012059465982019901
step: 130, loss: 0.014075016602873802
step: 140, loss: 0.08823948353528976
```

```
step: 150, loss: 0.07506641000509262
step: 160, loss: 0.22635413706302643
step: 170, loss: 0.008386767469346523
step: 180, loss: 0.010145291686058044
step: 190, loss: 0.15045902132987976
step: 200, loss: 0.01930258795619011
step: 210, loss: 0.018324226140975952
/usr/local/lib/python3.7/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=0.5647058823529412, f1=0.5462184873949579,
best_f1=0.5987525987525987
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03346949443221092
step: 10, loss: 0.017359759658575058
step: 20, loss: 0.035569895058870316
step: 30, loss: 0.027720734477043152
step: 40, loss: 0.0407809317111969
step: 50, loss: 0.07108347117900848
step: 60, loss: 0.06944723427295685
step: 70, loss: 0.036899179220199585
step: 80, loss: 0.011363442987203598
step: 90, loss: 0.0077889771200716496
step: 100, loss: 0.00251818192191422
step: 110, loss: 0.06579025834798813
step: 120, loss: 0.04196728393435478
step: 130, loss: 0.007788398768752813
step: 140, loss: 0.0554291196167469
step: 150, loss: 0.09418293833732605
step: 160, loss: 0.04895598441362381
step: 170, loss: 0.009784579277038574
step: 180, loss: 0.0005333093577064574
step: 190, loss: 0.018574541434645653
step: 200, loss: 0.01058240421116352
step: 210, loss: 0.05573860555887222
/usr/local/lib/python3.7/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=0.5797101449275361, f1=0.5308924485125858,
best_f1=0.5987525987525987
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18530310690402985
step: 10, loss: 0.0040881093591451645
step: 20, loss: 0.009964929893612862
```

```
step: 30, loss: 0.0027901914436370134
step: 40, loss: 0.10317505896091461
step: 50, loss: 0.0020893076434731483
step: 60, loss: 0.3020593523979187
step: 70, loss: 0.004532777704298496
step: 80, loss: 0.04394534230232239
step: 90, loss: 0.01608571968972683
step: 100, loss: 0.009812151081860065
step: 110, loss: 0.002480241237208247
step: 120, loss: 0.0012405210873112082
step: 130, loss: 0.016802005469799042
step: 140, loss: 0.006852015387266874
step: 150, loss: 0.05196940526366234
step: 160, loss: 0.004521131981164217
step: 170, loss: 0.001872727763839066
step: 180, loss: 0.08487411588430405
step: 190, loss: 0.037834007292985916
step: 200, loss: 0.05106090381741524
step: 210, loss: 0.062018394470214844
/usr/local/lib/python3.7/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=0.578544061302682, f1=0.5404255319148936,
best_f1=0.5987525987525987
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00445215031504631
step: 10, loss: 0.03087499551475048
step: 20, loss: 0.00359169440343976
step: 30, loss: 0.10981617122888565
step: 40, loss: 0.02248326689004898
step: 50, loss: 0.008056302554905415
step: 60, loss: 0.012057173065841198
step: 70, loss: 0.043218743056058884
step: 80, loss: 0.013420295901596546
step: 90, loss: 0.07126454263925552
step: 100, loss: 0.018655207008123398
step: 110, loss: 0.00041750731179490685
step: 120, loss: 0.01587611623108387
step: 130, loss: 0.07903938740491867
step: 140, loss: 0.009571732021868229
step: 150, loss: 0.0168557558208704
step: 160, loss: 0.0009462057496421039
step: 170, loss: 0.011119997128844261
step: 180, loss: 0.029031902551651
step: 190, loss: 0.08149389922618866
step: 200, loss: 0.021928848698735237
```

```
step: 210, loss: 0.036071620881557465
/usr/local/lib/python3.7/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=0.5903614457831325, f1=0.538293216630197,
best_f1=0.5987525987525987
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0033871782943606377
step: 10, loss: 0.0017875843914225698
step: 20, loss: 0.016724593937397003
step: 30, loss: 0.008433116599917412
step: 40, loss: 0.02964743971824646
step: 50, loss: 8.571975922677666e-05
step: 60, loss: 0.0012398340040817857
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 70, loss: 0.006941239815205336
step: 80, loss: 0.0001010161213343963
step: 90, loss: 0.033484071493148804
step: 100, loss: 9.994854917749763e-05
step: 110, loss: 0.0004712631634902209
step: 120, loss: 0.0006579289329238236
step: 130, loss: 0.002022832166403532
step: 140, loss: 0.0029154163785278797
step: 150, loss: 0.11009696871042252
step: 160, loss: 0.003153408644720912
step: 170, loss: 0.01997276395559311
step: 180, loss: 0.0012031731894239783
step: 190, loss: 0.01442489679902792
step: 200, loss: 0.01088443212211132
step: 210, loss: 0.001722204964607954
/usr/local/lib/python3.7/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=0.5613305613305614, f1=0.5300668151447662,
best_f1=0.5987525987525987
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007250833441503346
step: 10, loss: 0.001038626884110272
step: 20, loss: 0.00040666875429451466
step: 30, loss: 0.002915996592491865
step: 40, loss: 0.006604093592613935
step: 50, loss: 0.0012344716815277934
step: 60, loss: 0.000993320601992309
step: 70, loss: 0.0007152181351557374
```

```
step: 80, loss: 0.09951577335596085
step: 90, loss: 0.003765253582969308
step: 100, loss: 0.003527653869241476
step: 110, loss: 0.00068355501408204436
step: 120, loss: 0.00026242865715175867
step: 130, loss: 0.022024057805538177
step: 140, loss: 0.040662601590156555
step: 150, loss: 0.002396028721705079
step: 160, loss: 0.007715716026723385
step: 170, loss: 0.028905007988214493
step: 180, loss: 0.0019812986720353365
step: 190, loss: 0.0017809228738769889
step: 200, loss: 0.00025337544502690434
step: 210, loss: 0.00011713043204508722
/usr/local/lib/python3.7/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=0.576271186440678, f1=0.5270588235294117,
best_f1=0.5987525987525987
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002148967469111085
step: 10, loss: 0.000885773915797472
step: 20, loss: 0.002566723618656397
step: 30, loss: 0.00012256171612534672
step: 40, loss: 0.19232511520385742
step: 50, loss: 0.006171287503093481
step: 60, loss: 0.01236092671751976
step: 70, loss: 0.00104320771060884
step: 80, loss: 0.045744869858026505
step: 90, loss: 0.0074515691958367825
step: 100, loss: 0.003558909520506859
step: 110, loss: 0.0016215128125622869
step: 120, loss: 0.0021681934595108032
step: 130, loss: 0.0009909718064591289
step: 140, loss: 0.003923905082046986
step: 150, loss: 0.0002546290634199977
step: 160, loss: 0.0010671928757801652
step: 170, loss: 0.0004821609181817621
step: 180, loss: 0.000410845794249326
step: 190, loss: 0.038525309413671494
step: 200, loss: 0.0005013306508772075
step: 210, loss: 0.000901427585631609
/usr/local/lib/python3.7/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=0.5627705627705628, f1=0.5024390243902439,
```

```
best_f1=0.5987525987525987
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017011594027280807
step: 10, loss: 0.005293252412229776
step: 20, loss: 0.0008983829175122082
step: 30, loss: 0.01676170900464058
step: 40, loss: 0.00017770739214029163
step: 50, loss: 0.0033748208079487085
step: 60, loss: 0.036691613495349884
step: 70, loss: 0.06438285857439041
step: 80, loss: 0.0016931516584008932
step: 90, loss: 0.001092036021873355
step: 100, loss: 0.009816855192184448
step: 110, loss: 0.0004139520169701427
step: 120, loss: 0.0002338165504625067
step: 130, loss: 0.0001922969240695238
step: 140, loss: 0.013601619750261307
step: 150, loss: 0.00019246028386987746
step: 160, loss: 0.0008511441992595792
step: 170, loss: 0.005288091488182545
step: 180, loss: 0.001155282836407423
step: 190, loss: 0.00015932699898257852
step: 200, loss: 0.027109257876873016
step: 210, loss: 0.06894218176603317
/usr/local/lib/python3.7/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=0.5581395348837209, f1=0.49872773536895676,
best_f1=0.5987525987525987
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009396516834385693
step: 10, loss: 0.00013050866255071014
step: 20, loss: 0.0004387518856674433
step: 30, loss: 0.0019128072308376431
step: 40, loss: 0.00034186593256890774
step: 50, loss: 0.001202749670483172
step: 60, loss: 0.008149700239300728
step: 70, loss: 0.0073544918559491634
step: 80, loss: 0.059099976363182068
step: 90, loss: 0.003022068878635764
step: 100, loss: 0.0008468425367027521
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 110, loss: 0.0003831623762380332
step: 120, loss: 0.002323112217709422
```

```
step: 130, loss: 0.0005135294049978256
step: 140, loss: 0.0003106513759121299
step: 150, loss: 0.006269057746976614
step: 160, loss: 0.0002950870548374951
step: 170, loss: 0.004734999965876341
step: 180, loss: 0.0002642266917973757
step: 190, loss: 0.00015631128917448223
step: 200, loss: 0.00018501684826333076
step: 210, loss: 0.00030385481659322977
/usr/local/lib/python3.7/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=0.5789473684210527, f1=0.5530973451327434,
best_f1=0.5987525987525987
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008821626652032137
step: 10, loss: 0.014866198413074017
step: 20, loss: 0.0002316842001164332
step: 30, loss: 0.0020783967338502407
step: 40, loss: 0.00010839076276170009
step: 50, loss: 0.00035317783476784825
step: 60, loss: 0.0006632811855524778
step: 70, loss: 0.0002356784971198067
step: 80, loss: 0.0002977323019877076
step: 90, loss: 0.000561293214559555
step: 100, loss: 0.00019698776304721832
step: 110, loss: 0.0004001377383247018
step: 120, loss: 0.0003496749559417367
step: 130, loss: 0.0005965828895568848
step: 140, loss: 0.0020746877416968346
step: 150, loss: 0.0003310703032184392
step: 160, loss: 0.02401682361960411
step: 170, loss: 0.00022006593644618988
step: 180, loss: 0.0021635883022099733
step: 190, loss: 0.0014808080159127712
step: 200, loss: 0.0006741896504536271
step: 210, loss: 0.032473139464855194
/usr/local/lib/python3.7/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=0.5637860082304527, f1=0.5495495495495496,
best_f1=0.5987525987525987
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.")
2293it [00:04, 559.81it/s]
load_f1 = 0.5684647302904564
real_f1 = 0.5785123966942148
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 [00:12, 363.35it/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 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.

298

```
  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: 0.8526515364646912
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.1622495949268341
step: 20, loss: 0.15290598571300507
step: 30, loss: 0.5114005208015442
step: 40, loss: 0.2606736123561859
step: 50, loss: 0.30762121081352234
step: 60, loss: 0.3663616180419922
step: 70, loss: 0.17829003930091858
step: 80, loss: 0.5369257926940918
step: 90, loss: 0.25390905141830444
step: 100, loss: 0.2227405607700348
step: 110, loss: 0.23586991429328918
step: 120, loss: 0.4210648536682129
step: 130, loss: 0.3447793126106262
step: 140, loss: 0.3256511092185974
step: 150, loss: 0.27603963017463684
step: 160, loss: 0.2025219351053238
step: 170, loss: 0.3663181662559509
step: 180, loss: 0.24232590198516846
step: 190, loss: 0.12980054318904877
/usr/local/lib/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=0.5977653631284916, f1=0.6050420168067228,
best_f1=0.6050420168067228
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22265268862247467
step: 10, loss: 0.040102552622556686
step: 20, loss: 0.05247143283486366
step: 30, loss: 0.11946551501750946
```

```
step: 40, loss: 0.44515594840049744
step: 50, loss: 0.3054981231689453
step: 60, loss: 0.14509043097496033
step: 70, loss: 0.25588637590408325
step: 80, loss: 0.19670814275741577
step: 90, loss: 0.11994846165180206
step: 100, loss: 0.1868390440940857
step: 110, loss: 0.15938939154148102
step: 120, loss: 0.1835823655128479
step: 130, loss: 0.26637905836105347
step: 140, loss: 0.24194839596748352
step: 150, loss: 0.03741403669118881
step: 160, loss: 0.014675416983664036
step: 170, loss: 0.1766241490840912
step: 180, loss: 0.10890578478574753
step: 190, loss: 0.13231849670410156
/usr/local/lib/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=0.7769028871391076, f1=0.774025974025974,
best_f1=0.774025974025974
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10239367187023163
step: 10, loss: 0.17543761432170868
step: 20, loss: 0.07858606427907944
step: 30, loss: 0.03926929458975792
step: 40, loss: 0.05929631367325783
step: 50, loss: 0.04984656721353531
step: 60, loss: 0.08101975172758102
step: 70, loss: 0.16007143259048462
step: 80, loss: 0.08586083352565765
step: 90, loss: 0.05705208331346512
step: 100, loss: 0.10723020136356354
step: 110, loss: 0.1687435358762741
step: 120, loss: 0.017839059233665466
step: 130, loss: 0.07937224209308624
step: 140, loss: 0.14172323048114777
step: 150, loss: 0.1412457972764969
step: 160, loss: 0.04473868012428284
step: 170, loss: 0.039912715554237366
step: 180, loss: 0.07240824401378632
step: 190, loss: 0.09782250225543976
```

```
/usr/local/lib/python3.7/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=0.7513812154696132, f1=0.7603305785123966,
best_f1=0.774025974025974
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0888356938958168
step: 10, loss: 0.02912314422428608
step: 20, loss: 0.0212752353399992
step: 30, loss: 0.05465702712535858
step: 40, loss: 0.004202126991003752
step: 50, loss: 0.044260650873184204
step: 60, loss: 0.10900484025478363
step: 70, loss: 0.03122742474079132
step: 80, loss: 0.015109182335436344
step: 90, loss: 0.11619798094034195
step: 100, loss: 0.008591465651988983
step: 110, loss: 0.05849190801382065
step: 120, loss: 0.051339101046323776
step: 130, loss: 0.22791963815689087
step: 140, loss: 0.03420155495405197
step: 150, loss: 0.03929696977138519
step: 160, loss: 0.11529029905796051
step: 170, loss: 0.010199062526226044
step: 180, loss: 0.338693767786026
step: 190, loss: 0.04471869766712189
/usr/local/lib/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=0.7864583333333334, f1=0.7757255936675462,
best_f1=0.7757255936675462
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025892900303006172
step: 10, loss: 0.006460359785705805
step: 20, loss: 0.14982633292675018
step: 30, loss: 0.015936024487018585
step: 40, loss: 0.07890917360782623
step: 50, loss: 0.07486417889595032
step: 60, loss: 0.06983957439661026
step: 70, loss: 0.001751126372255385
```

```
step: 80, loss: 0.04699936881661415
step: 90, loss: 0.0355713814496994
step: 100, loss: 0.13152584433555603
step: 110, loss: 0.07234766334295273
step: 120, loss: 0.0036161586176604033
step: 130, loss: 0.005376610439270735
step: 140, loss: 0.062213234603405
step: 150, loss: 0.07407143712043762
step: 160, loss: 0.053862396627664566
step: 170, loss: 0.020088912919163704
step: 180, loss: 0.01716466434299946
step: 190, loss: 0.1432359516620636
/usr/local/lib/python3.7/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=0.7658402203856748, f1=0.7578347578347578,
best_f1=0.7757255936675462
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.048325277864933014
step: 10, loss: 0.0007888911641202867
step: 20, loss: 0.0046721515245735645
step: 30, loss: 0.0025475251022726297
step: 40, loss: 0.011242321692407131
step: 50, loss: 0.011584515683352947
step: 60, loss: 0.007028626278042793
step: 70, loss: 0.10061197727918625
step: 80, loss: 0.008562526665627956
step: 90, loss: 0.012151284143328667
step: 100, loss: 0.004281593952327967
step: 110, loss: 0.0020138765685260296
step: 120, loss: 0.0023856598418205976
step: 130, loss: 0.0006609336123801768
step: 140, loss: 0.0521550290286541
step: 150, loss: 0.01702267676591873
step: 160, loss: 0.032913099974393845
step: 170, loss: 0.03399580717086792
step: 180, loss: 0.0023857702035456896
step: 190, loss: 0.0513668917119503
/usr/local/lib/python3.7/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=0.7513513513513514, f1=0.7272727272727273,
best_f1=0.7757255936675462
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.03304577246308327
step: 10, loss: 0.005141180474311113
step: 20, loss: 0.0014274833956733346
step: 30, loss: 0.022616252303123474
step: 40, loss: 0.006060120649635792
step: 50, loss: 0.0053943139500916
step: 60, loss: 0.0024392781779170036
step: 70, loss: 0.060688961297273636
step: 80, loss: 0.014590129256248474
step: 90, loss: 0.002208179095759988
step: 100, loss: 0.0029669920913875103
step: 110, loss: 0.02777290530502796
step: 120, loss: 0.0013899190817028284
step: 130, loss: 0.005412947852164507
step: 140, loss: 0.00039452247438021004
step: 150, loss: 0.005737514700740576
step: 160, loss: 0.0003618115733843297
step: 170, loss: 0.006106486544013023
step: 180, loss: 0.0935111716389656
step: 190, loss: 0.0015339636011049151
/usr/local/lib/python3.7/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=0.7567567567567567, f1=0.7606382978723404,
best_f1=0.7757255936675462
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005399833898991346
step: 10, loss: 0.004758899565786123
step: 20, loss: 0.0006902448949404061
step: 30, loss: 0.021595532074570656
step: 40, loss: 0.014070981182157993
step: 50, loss: 0.0011844885302707553
step: 60, loss: 0.0003598544280976057
step: 70, loss: 0.0011272775009274483
step: 80, loss: 0.007725153584033251
step: 90, loss: 0.0016259347321465611
step: 100, loss: 0.0078541599214077
step: 110, loss: 0.001069218385964632
step: 120, loss: 0.0003348469326738268
step: 130, loss: 0.010573204606771469
step: 140, loss: 0.0026080841198563576
step: 150, loss: 0.0176596287637949
step: 160, loss: 0.03772908076643944
step: 170, loss: 0.13740362226963043
step: 180, loss: 0.004652864765375853
step: 190, loss: 0.0831577256321907
```

```
/usr/local/lib/python3.7/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=0.7379679144385027, f1=0.7345844504021448,
best_f1=0.7757255936675462
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007003562059253454
step: 10, loss: 0.012800557538866997
step: 20, loss: 0.004343134816735983
step: 30, loss: 0.028703883290290833
step: 40, loss: 0.010633555240929127
step: 50, loss: 0.04496218264102936
step: 60, loss: 0.0021449087653309107
step: 70, loss: 0.1128157526254654
step: 80, loss: 0.0009956260910257697
step: 90, loss: 0.0031406471971422434
step: 100, loss: 0.0050194947980344296
step: 110, loss: 0.005283646751195192
step: 120, loss: 0.12878990173339844
step: 130, loss: 0.0018736198544502258
step: 140, loss: 0.02231176197528839
step: 150, loss: 0.0027489231433719397
step: 160, loss: 0.00096565514104440081
step: 170, loss: 0.0009099089074879885
step: 180, loss: 0.0034272028133273125
step: 190, loss: 0.0005584045429714024
/usr/local/lib/python3.7/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=0.7789473684210525, f1=0.7376623376623376,
best_f1=0.7757255936675462
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0063545010052621365
step: 10, loss: 0.0003115233557764441
step: 20, loss: 0.0029619038105010986
step: 30, loss: 0.17283298075199127
step: 40, loss: 0.00014169128553476185
step: 50, loss: 0.00018542149337008595
step: 60, loss: 0.011540194042026997
step: 70, loss: 0.003310675732791424
step: 80, loss: 0.00034451045212335885
step: 90, loss: 0.004304507281631231
step: 100, loss: 0.0003441321023274213
step: 110, loss: 0.0006000292487442493
```

```
step: 120, loss: 0.0031111373100429773
step: 130, loss: 0.002161916811019182
step: 140, loss: 0.00045840759412385523
step: 150, loss: 0.00023131629859562963
step: 160, loss: 0.0015755644999444485
step: 170, loss: 0.00042498260154388845
step: 180, loss: 0.0339776873588562
step: 190, loss: 0.012339742854237556
/usr/local/lib/python3.7/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=0.7557840616966581, f1=0.7187499999999999,
best_f1=0.7757255936675462
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015822557732462883
step: 10, loss: 0.000632092822343111
step: 20, loss: 0.003214964410290122
step: 30, loss: 0.0010680697159841657
step: 40, loss: 0.00037067331140860915
step: 50, loss: 0.0002281337947351858
step: 60, loss: 0.0005243667983449996
step: 70, loss: 0.005239700898528099
step: 80, loss: 0.0013722919393330812
step: 90, loss: 0.0004165449063293636
step: 100, loss: 0.0001772027026163414
step: 110, loss: 0.00016900508489925414
step: 120, loss: 0.000242379741393961
step: 130, loss: 0.0006673146272078156
step: 140, loss: 0.0005868463777005672
step: 150, loss: 0.016672244295477867
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 160, loss: 0.0002286597591591999
step: 170, loss: 0.0021058714482933283
step: 180, loss: 0.004739684984087944
step: 190, loss: 0.00022533519950229675
/usr/local/lib/python3.7/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=0.7547169811320755, f1=0.7479224376731302,
best_f1=0.7757255936675462
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017460146045777947
step: 10, loss: 0.002183436183258891
step: 20, loss: 0.0009300124365836382
```

```
step: 30, loss: 0.0022175032645463943
step: 40, loss: 0.007089582271873951
step: 50, loss: 0.0013313277158886194
step: 60, loss: 0.0007274263771250844
step: 70, loss: 0.0008528248290531337
step: 80, loss: 0.0002710193512029946
step: 90, loss: 0.001095544663257897
step: 100, loss: 0.00041028836858458817
step: 110, loss: 0.00027864283765666187
step: 120, loss: 0.00018908196943812072
step: 130, loss: 0.00020076178770978004
step: 140, loss: 0.0028515516314655542
step: 150, loss: 0.00020262632460799068
step: 160, loss: 0.0002656590659171343
step: 170, loss: 0.00021174144058022648
step: 180, loss: 0.0006842211005277932
step: 190, loss: 0.0024566075298935175
/usr/local/lib/python3.7/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=0.7548209366391184, f1=0.7415730337078652,
best_f1=0.7757255936675462
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004400421748869121
step: 10, loss: 0.0005130230565555394
step: 20, loss: 0.0012556889560073614
step: 30, loss: 0.000607928610406816
step: 40, loss: 0.00016856241563800722
step: 50, loss: 0.0003315576759632677
step: 60, loss: 0.0019781216979026794
step: 70, loss: 0.0003189630515407771
step: 80, loss: 0.0032603370491415262
step: 90, loss: 0.00023574582883156836
step: 100, loss: 0.0006554744904860854
step: 110, loss: 0.0005859971279278398
step: 120, loss: 0.004890565760433674
step: 130, loss: 0.00042906717862933874
step: 140, loss: 0.0009230130235664546
step: 150, loss: 0.00019966256513725966
step: 160, loss: 0.00029691733652725816
step: 170, loss: 0.00016496233001817018
step: 180, loss: 0.006165761034935713
step: 190, loss: 0.002340832492336631
/usr/local/lib/python3.7/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=0.7526881720430108, f1=0.7298050139275766,
best_f1=0.7757255936675462
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013457302702590823
step: 10, loss: 0.0005029713502153754
step: 20, loss: 0.010156230069696903
step: 30, loss: 0.0022251259069889784
step: 40, loss: 0.0005805251421406865
step: 50, loss: 0.001362494076602161
step: 60, loss: 0.00022929938859306276
step: 70, loss: 0.0005194447585381567
step: 80, loss: 0.0004024539957754314
step: 90, loss: 0.013273858465254307
step: 100, loss: 0.00038078747456893325
step: 110, loss: 0.0011311584385111928
step: 120, loss: 0.0005576945259235799
step: 130, loss: 0.00014079395623411983
step: 140, loss: 0.0007099188514985144
step: 150, loss: 0.0002705720253288746
step: 160, loss: 0.0009576361044310033
step: 170, loss: 0.0002073946816381067
step: 180, loss: 0.0010190039174631238
step: 190, loss: 0.00043159068445675075
/usr/local/lib/python3.7/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=0.7567567567567567, f1=0.7409470752089136,
best_f1=0.7757255936675462
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011838052421808243
step: 10, loss: 0.00017929247405845672
step: 20, loss: 0.000500953639857471
step: 30, loss: 0.00037309847539290786
step: 40, loss: 0.00079893390648067
step: 50, loss: 0.00047013425501063466
step: 60, loss: 0.00031858228612691164
step: 70, loss: 0.003491707146167755
step: 80, loss: 0.0005452118348330259
step: 90, loss: 0.0012996392324566841
step: 100, loss: 0.001707019517198205
step: 110, loss: 0.0002213877742178738
step: 120, loss: 0.0006461346056312323
step: 130, loss: 0.0003944840282201767
step: 140, loss: 0.002853600075468421
```

```
step: 150, loss: 0.0005710538243874907
step: 160, loss: 0.0010142637183889747
step: 170, loss: 0.0002433463669149205
step: 180, loss: 0.00045148676144890487
step: 190, loss: 0.009301840327680111
/usr/local/lib/python3.7/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=0.7566137566137565, f1=0.7377049180327869,
best_f1=0.7757255936675462
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.")
2049it [00:07, 284.70it/s]
load_f1 = 0.6951871657754011
real_f1 = 0.6805194805194804
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 [00:13, 330.81it/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: 0.8403683304786682
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.2178473174571991
step: 20, loss: 0.1599653661251068
step: 30, loss: 0.2423805594444275
step: 40, loss: 0.3089097738265991
step: 50, loss: 0.3747328519821167
step: 60, loss: 0.44012370705604553
step: 70, loss: 0.3085483908653259
step: 80, loss: 0.2549421787261963
step: 90, loss: 0.4069548547267914
step: 100, loss: 0.23619809746742249
step: 110, loss: 0.18247516453266144
step: 120, loss: 0.5370160937309265
step: 130, loss: 0.42224031686782837
step: 140, loss: 0.46858471632003784
step: 150, loss: 0.12686139345169067
step: 160, loss: 0.2710825502872467
step: 170, loss: 0.14500205218791962
/usr/local/lib/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=0.6157517899761337, f1=0.6157635467980296,
best_f1=0.6157635467980296
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.258387953042984
step: 10, loss: 0.05519657954573631
step: 20, loss: 0.2603580057621002
step: 30, loss: 0.20256413519382477
step: 40, loss: 0.19206485152244568
step: 50, loss: 0.20584805309772491
step: 60, loss: 0.14003658294677734
step: 70, loss: 0.2197514921426773
step: 80, loss: 0.1394728273153305
step: 90, loss: 0.14538191258907318
step: 100, loss: 0.09314236044883728
step: 110, loss: 0.22254718840122223
step: 120, loss: 0.043830521404743195
step: 130, loss: 0.05650699511170387
step: 140, loss: 0.08738183230161667
step: 150, loss: 0.32008063793182373
step: 160, loss: 0.1636134833097458
step: 170, loss: 0.1106729805469513
/usr/local/lib/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=0.7333333333333333, f1=0.744920993227991,
best_f1=0.744920993227991
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09714311361312866
step: 10, loss: 0.031635116785764694
step: 20, loss: 0.06750050187110901
step: 30, loss: 0.3009938895702362
step: 40, loss: 0.012129572220146656
step: 50, loss: 0.051496926695108414
step: 60, loss: 0.19935691356658936
step: 70, loss: 0.12255066633224487
step: 80, loss: 0.27661874890327454
step: 90, loss: 0.09973596781492233

```
step: 100, loss: 0.01913340575993061
step: 110, loss: 0.07184303551912308
step: 120, loss: 0.0070182993076741695
step: 130, loss: 0.1402713507413864
step: 140, loss: 0.009591648355126381
step: 150, loss: 0.08518731594085693
step: 160, loss: 0.0454288087785244
step: 170, loss: 0.13902407884597778
/usr/local/lib/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=0.7473460721868365, f1=0.717741935483871,
best_f1=0.717741935483871
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03194158524274826
step: 10, loss: 0.11421363055706024
step: 20, loss: 0.04057404026389122
step: 30, loss: 0.018170922994613647
step: 40, loss: 0.13086000084877014
step: 50, loss: 0.006035767495632172
step: 60, loss: 0.07954858988523483
step: 70, loss: 0.0036943508312106133
step: 80, loss: 0.03763582184910774
step: 90, loss: 0.0346662811934948
step: 100, loss: 0.01589970290660858
step: 110, loss: 0.06445785611867905
step: 120, loss: 0.13243988156318665
step: 130, loss: 0.07756944745779037
step: 140, loss: 0.016376616433262825
step: 150, loss: 0.044153761118650436
step: 160, loss: 0.060139499604701996
step: 170, loss: 0.14397849142551422
/usr/local/lib/python3.7/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=0.7250608272506082, f1=0.7433628318584071,
best_f1=0.717741935483871
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09246410429477692
step: 10, loss: 0.08496850728988647
```

```
step: 20, loss: 0.011437688022851944
step: 30, loss: 0.022500917315483093
step: 40, loss: 0.0035146998707205057
step: 50, loss: 0.006990667898207903
step: 60, loss: 0.0016755693359300494
step: 70, loss: 0.051847901195287704
step: 80, loss: 0.024034764617681503
step: 90, loss: 0.06698314845561981
step: 100, loss: 0.012990486808121204
step: 110, loss: 0.1807357221841812
step: 120, loss: 0.0460970439016819
step: 130, loss: 0.0037277634255588055
step: 140, loss: 0.04603267088532448
step: 150, loss: 0.007096074055880308
step: 160, loss: 0.006410597357898951
step: 170, loss: 0.04705102741718292
/usr/local/lib/python3.7/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=0.7358024691358026, f1=0.7517401392111369,
best_f1=0.717741935483871
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016402633860707283
step: 10, loss: 0.01728544384241104
step: 20, loss: 0.00371820037253201
step: 30, loss: 0.04639582335948944
step: 40, loss: 0.09103789180517197
step: 50, loss: 0.011905331164598465
step: 60, loss: 0.10286680608987808
step: 70, loss: 0.007487718015909195
step: 80, loss: 0.01425759494304657
step: 90, loss: 0.004731804132461548
step: 100, loss: 0.005550788249820471
step: 110, loss: 0.023885153234004974
step: 120, loss: 0.07462793588638306
step: 130, loss: 0.22476644814014435
step: 140, loss: 0.1470572054386139
step: 150, loss: 0.13786432147026062
step: 160, loss: 0.04325150325894356
step: 170, loss: 0.013548823073506355
/usr/local/lib/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=0.7719298245614035, f1=0.7645687645687645,
best_f1=0.7645687645687645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004432481713593006
step: 10, loss: 0.0005845482228323817
step: 20, loss: 0.007190740201622248
step: 30, loss: 0.07359065115451813
step: 40, loss: 0.014121891930699348
step: 50, loss: 0.02755906991660595
step: 60, loss: 0.1659705936908722
step: 70, loss: 0.0027587967924773693
step: 80, loss: 0.05014418438076973
step: 90, loss: 0.01770656742155552
step: 100, loss: 0.004358977545052767
step: 110, loss: 0.0069578420370817184
step: 120, loss: 0.14743220806121826
step: 130, loss: 0.03144379332661629
step: 140, loss: 0.002918673912063241
step: 150, loss: 0.1031666025519371
step: 160, loss: 0.007205810863524675
step: 170, loss: 0.02908414974808693
/usr/local/lib/python3.7/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=0.755784061696658, f1=0.7625899280575539,
best_f1=0.7645687645687645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03001733496785164
step: 10, loss: 0.0016001465264707804
step: 20, loss: 0.007002921309322119
step: 30, loss: 0.027753649279475212
step: 40, loss: 0.00044897443149238825
step: 50, loss: 0.0008892653277143836
step: 60, loss: 0.0026750629767775536
step: 70, loss: 0.02505616471171379
step: 80, loss: 0.003657606430351734
step: 90, loss: 0.004410834517329931
step: 100, loss: 0.0027569078374654055
step: 110, loss: 0.06839301437139511
step: 120, loss: 0.0005193359102122486
step: 130, loss: 0.00151583191473037
step: 140, loss: 0.0019804888870567083
step: 150, loss: 0.019713472574949265
```

```
step: 160, loss: 0.023904841393232346
step: 170, loss: 0.00459772115573287
/usr/local/lib/python3.7/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=0.7523364485981308, f1=0.7379912663755459,
best_f1=0.7645687645687645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017227388452738523
step: 10, loss: 0.0021298956125974655
step: 20, loss: 0.0006127889500930905
step: 30, loss: 0.006445101462304592
step: 40, loss: 0.012043917551636696
step: 50, loss: 0.02635873854160309
step: 60, loss: 0.025315331295132637
step: 70, loss: 0.010288802906870842
step: 80, loss: 0.00727156363427639
step: 90, loss: 0.011466717347502708
step: 100, loss: 0.023057514801621437
step: 110, loss: 0.00030678685288876295
step: 120, loss: 0.04030816629528999
step: 130, loss: 0.0003082749026361853
step: 140, loss: 0.036168161779642105
step: 150, loss: 0.025264566764235497
step: 160, loss: 0.014862197451293468
step: 170, loss: 0.023590072989463806
/usr/local/lib/python3.7/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=0.7322654462242564, f1=0.7061310782241015,
best_f1=0.7645687645687645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.026106031611561775
step: 10, loss: 0.02099969983100891
step: 20, loss: 0.0011449063895270228
step: 30, loss: 0.00021806808945257217
step: 40, loss: 0.0011534603545442224
step: 50, loss: 0.0033233079593628645
step: 60, loss: 0.0031335176900029182
step: 70, loss: 0.0033368209842592478
step: 80, loss: 0.0027243997901678085
step: 90, loss: 0.0005161263979971409
step: 100, loss: 0.0004963356186635792
step: 110, loss: 0.0008272267878055573
```

```
step: 120, loss: 0.002906476380303502
step: 130, loss: 0.019617201760411263
step: 140, loss: 0.030093956738710403
step: 150, loss: 0.00041213969234377146
step: 160, loss: 0.006398243363946676
step: 170, loss: 0.0015398883260786533
/usr/local/lib/python3.7/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=0.7226463104325701, f1=0.7535545023696684,
best_f1=0.7645687645687645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00041461965884082020
step: 10, loss: 0.00114936521276883163
step: 20, loss: 0.010272394865751266
step: 30, loss: 0.0009224284440279007
step: 40, loss: 0.06864377111196518
step: 50, loss: 0.0015520559391006827
step: 60, loss: 0.0022198997903615236
step: 70, loss: 0.0015357087831944227
step: 80, loss: 0.004930688533931971
step: 90, loss: 0.0002941093989647925
step: 100, loss: 0.004856475628912449
step: 110, loss: 0.0004943090607412159
step: 120, loss: 0.004947804380208254
step: 130, loss: 0.00859283097088337
step: 140, loss: 0.025194162502884865
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 150, loss: 0.03193634748458862
step: 160, loss: 0.0009383437572978437
step: 170, loss: 0.12437556684017181
/usr/local/lib/python3.7/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=0.7669172932330827,
best_f1=0.7645687645687645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015228138072416186
step: 10, loss: 0.003220732556656003
step: 20, loss: 0.003156794933602214
step: 30, loss: 0.0025534399319440126
step: 40, loss: 0.0003448036150075495
step: 50, loss: 0.0012647476978600025
step: 60, loss: 0.0003665366966743022
```

```
step: 70, loss: 0.004279526881873608
step: 80, loss: 0.011164015159010887
step: 90, loss: 0.0003750711039174348
step: 100, loss: 0.001297448412515223
step: 110, loss: 0.0002350252034375444
step: 120, loss: 0.0007683291332796216
step: 130, loss: 0.0009205978712998331
step: 140, loss: 0.0003248266875743866
step: 150, loss: 0.0009638978517614305
step: 160, loss: 0.00022662850096821785
step: 170, loss: 0.0033537549898028374
/usr/local/lib/python3.7/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=0.7035175879396984, f1=0.7319347319347319,
best_f1=0.7645687645687645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002507115714251995
step: 10, loss: 0.02715686894953251
step: 20, loss: 0.001729228999465704
step: 30, loss: 0.00020280040916986763
step: 40, loss: 0.0002670518006198108
step: 50, loss: 0.0004491469298955053
step: 60, loss: 0.00018277709023095667
step: 70, loss: 0.00089611177663877606
step: 80, loss: 0.00026853810413740575
step: 90, loss: 0.002974635921418667
step: 100, loss: 0.00021529558580368757
step: 110, loss: 0.0013072615256533027
step: 120, loss: 0.00058805245600640774
step: 130, loss: 0.0023012831807136536
step: 140, loss: 0.00014370432472787797
step: 150, loss: 0.0010053126607090235
step: 160, loss: 0.0020146903116256
step: 170, loss: 0.0005083572468720376
/usr/local/lib/python3.7/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=0.7029702970297029, f1=0.7339449541284403,
best_f1=0.7645687645687645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005254154675640166
step: 10, loss: 0.0006337572121992707
step: 20, loss: 0.012205451726913452
```

```
step: 30, loss: 0.0002543435257393867
step: 40, loss: 0.0015890832291916013
step: 50, loss: 0.0036837481893599033
step: 60, loss: 0.00019017122394870967
step: 70, loss: 0.010181419551372528
step: 80, loss: 0.0009665945544838905
step: 90, loss: 0.0002089125191560015
step: 100, loss: 0.00023291418619919568
step: 110, loss: 0.0015163217904046178
step: 120, loss: 0.0003150603442918509
step: 130, loss: 0.0022665360011160374
step: 140, loss: 0.0008425480336882174
step: 150, loss: 0.0006912583485245705
step: 160, loss: 0.002464793622493744
step: 170, loss: 0.0002564960450399667
/usr/local/lib/python3.7/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=0.7183462532299741, f1=0.7616707616707616,
best_f1=0.7645687645687645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02021007612347603
step: 10, loss: 0.005965689662843943
step: 20, loss: 0.01310526393353939
step: 30, loss: 0.010721871629357338
step: 40, loss: 0.0005853326874785125
step: 50, loss: 0.00327649223618209364
step: 60, loss: 0.0023774844594299793
step: 70, loss: 0.08352729678153992
step: 80, loss: 0.0003729189920704812
step: 90, loss: 0.00012251715816091746
step: 100, loss: 0.0009406227036379278
step: 110, loss: 0.00023011973826214677
step: 120, loss: 0.0015363874845206738
step: 130, loss: 0.0024131585378199816
step: 140, loss: 0.0043959952890872955
step: 150, loss: 0.00559359323233366
step: 160, loss: 0.0001529937144368887
step: 170, loss: 0.00034058454912155867
/usr/local/lib/python3.7/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=0.7168831168831169, f1=0.7696078431372548,
best_f1=0.7645687645687645
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.")
1916it [00:04, 416.08it/s]
load_f1 = 0.595482546201232
real_f1 = 0.5742574257425742
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 [00:13, 324.78it/s]
```

3.3 DK DIRTY

3.3.1 DBLP-ACM - Running the matcher

```python
# 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 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: 0.8023170232772827
step: 10, loss: 0.4656188488006592
step: 20, loss: 0.5780977010726929
step: 30, loss: 0.43250858783721924
step: 40, loss: 0.17409218847751617
step: 50, loss: 0.19680823385715485
step: 60, loss: 0.22596286237239838
step: 70, loss: 0.20372311770915985
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 80, loss: 0.3168289065361023
step: 90, loss: 0.06525294482707977
step: 100, loss: 0.13796879351139069
step: 110, loss: 0.11110221594572067
step: 120, loss: 0.08606461435556412
step: 130, loss: 0.008893309161067009
step: 140, loss: 0.10000675171613693
step: 150, loss: 0.12347040325403214
step: 160, loss: 0.2510738670825958
step: 170, loss: 0.011755900457501411
step: 180, loss: 0.007814967073500156
step: 190, loss: 0.049143191426992416
step: 200, loss: 0.021973423659801483
step: 210, loss: 0.019783250987529755
step: 220, loss: 0.018482521176338196
step: 230, loss: 0.024550314992666245
/usr/local/lib/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=0.9537444933920706, f1=0.9595505617977529,

```
best_f1=0.9595505617977529
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16766905784606934
step: 10, loss: 0.1483897715806961
step: 20, loss: 0.014369934797286987
step: 30, loss: 0.0053037069737911224
step: 40, loss: 0.030080940574407578
step: 50, loss: 0.03980958089232445
step: 60, loss: 0.07678020000457764
step: 70, loss: 0.04247530177235603
step: 80, loss: 0.004264447372406721
step: 90, loss: 0.008468168787658215
step: 100, loss: 0.11982189863920212
step: 110, loss: 0.011264697648584843
step: 120, loss: 0.006137117277830839
step: 130, loss: 0.005555499345064163
step: 140, loss: 0.14937730133533478
step: 150, loss: 0.023834677413105965
step: 160, loss: 0.010567196644842625
step: 170, loss: 0.027004549279808998
step: 180, loss: 0.013722257688641548
step: 190, loss: 0.1922440081834793
step: 200, loss: 0.12258368730545044
step: 210, loss: 0.07370546460151672
step: 220, loss: 0.002871597185730934
step: 230, loss: 0.05371581390500069
/usr/local/lib/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=0.9658213891951488, f1=0.9655172413793103,
best_f1=0.9655172413793103
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18989478051662445
step: 10, loss: 0.015083400532603264
step: 20, loss: 0.020328039303421974
step: 30, loss: 0.007786072790622711
step: 40, loss: 0.0214958805590868
step: 50, loss: 0.028253242215722084
step: 60, loss: 0.24367040395736694
step: 70, loss: 0.0034396154806017876
```

```
step: 80, loss: 0.01778123900294304
step: 90, loss: 0.11374447494745255
step: 100, loss: 0.005425593815743923
step: 110, loss: 0.008776855655014515
step: 120, loss: 0.05409196391701698
step: 130, loss: 0.00210560648702085
step: 140, loss: 0.0013315471587702632
step: 150, loss: 0.0024322352837771177
step: 160, loss: 0.004644227679818869
step: 170, loss: 0.0016112611629068851
step: 180, loss: 0.009248118847608566
step: 190, loss: 0.13212499022483826
step: 200, loss: 0.05641099810600281
step: 210, loss: 0.02567894756793976
step: 220, loss: 0.024899888783693314
step: 230, loss: 0.13410641252994537
/usr/local/lib/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=0.9673790776152981, 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: 0.004758761264383793
step: 10, loss: 0.01597563922405243
step: 20, loss: 0.007742972578853369
step: 30, loss: 0.0011405862169340253
step: 40, loss: 0.007283931132405996
step: 50, loss: 0.000898194033652544
step: 60, loss: 0.001272450783289969
step: 70, loss: 0.011212275363504887
step: 80, loss: 0.2466096132993698
step: 90, loss: 0.09325653314590454
step: 100, loss: 0.0016982865054160357
step: 110, loss: 0.011958848685026169
step: 120, loss: 0.02208198606967926
step: 130, loss: 0.06767036765813828
step: 140, loss: 0.022285401821136475
step: 150, loss: 0.005727482959628105
step: 160, loss: 0.013670049607753754
step: 170, loss: 0.0007048522238619626
step: 180, loss: 0.10027561336755753
step: 190, loss: 0.0035584603901952505
```

```
step: 200, loss: 0.003259016200900078
step: 210, loss: 0.04887103661894798
step: 220, loss: 0.03475504368543625
step: 230, loss: 0.000466061464976151935
/usr/local/lib/python3.7/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=0.9657458563535912, f1=0.9656699889258028,
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: 0.0016885102959349751
step: 10, loss: 0.0019611367024481297
step: 20, loss: 0.0014248511288315058
step: 30, loss: 0.0012098532170057297
step: 40, loss: 0.004048175178468227
step: 50, loss: 0.00042719359043985605
step: 60, loss: 0.017000798135995865
step: 70, loss: 0.00026113720377907157
step: 80, loss: 0.02285776287317276
step: 90, loss: 0.003801458515226841
step: 100, loss: 0.005983484443277121
step: 110, loss: 0.0005145093891769648
step: 120, loss: 0.0638607069849968
step: 130, loss: 0.04753407835960388
step: 140, loss: 0.002770161023363471
step: 150, loss: 0.0015476770931854844
step: 160, loss: 0.035063695162534714
step: 170, loss: 0.09125825762748718
step: 180, loss: 0.0013715462991967797
step: 190, loss: 0.0014345372328534722
step: 200, loss: 0.0039053724613040686
step: 210, loss: 0.0008873804472386837
step: 220, loss: 0.06434004008769989
step: 230, loss: 0.00842159911990166
/usr/local/lib/python3.7/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=0.9668874172185431, f1=0.9658213891951488,
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: 0.0018108980730175972
step: 10, loss: 0.0022691134363412857
step: 20, loss: 0.0030016948003321886
step: 30, loss: 0.001127022784203291
```

```
step: 40, loss: 0.0035230012144893408
step: 50, loss: 0.010563163086771965
step: 60, loss: 0.0018260659417137504
step: 70, loss: 0.0011850178707391024
step: 80, loss: 0.0006442281301133335
step: 90, loss: 0.0022284602746367455
step: 100, loss: 0.0004938521306030452
step: 110, loss: 0.0106525924056776842
step: 120, loss: 0.0015418510884046555
step: 130, loss: 0.0007641388801857829
step: 140, loss: 0.0018768158042803407
step: 150, loss: 0.0014626665506511927
step: 160, loss: 0.0038092159666121006
step: 170, loss: 0.0009543675114400685
step: 180, loss: 0.000965174229349941
step: 190, loss: 0.013672786764800549
step: 200, loss: 0.009690510109066963
step: 210, loss: 0.000745739380363375
step: 220, loss: 0.00033360658562742174
step: 230, loss: 0.0003636429028119892
/usr/local/lib/python3.7/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=0.9624724061810155, f1=0.9547960308710033,
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: 0.17833131551742554
step: 10, loss: 0.0005280391196720302
step: 20, loss: 0.00032381757046096027
step: 30, loss: 0.0026095244102180004
step: 40, loss: 0.15599803626537323
step: 50, loss: 0.0003886458871420473
step: 60, loss: 0.02353602834045887
step: 70, loss: 0.000238978405832313
step: 80, loss: 0.0006670372094959021
step: 90, loss: 0.004841422196477652
step: 100, loss: 0.00029697915306314826
step: 110, loss: 0.0014577015535905957
step: 120, loss: 0.0005644787452183664
step: 130, loss: 0.0017113413196057081
step: 140, loss: 0.0001864821242634207
step: 150, loss: 0.1691366732120514
step: 160, loss: 0.00039563531754538417
step: 170, loss: 0.0005330440471880138
step: 180, loss: 0.0009419225971214473
step: 190, loss: 0.005254759453237057
```

```
step: 200, loss: 0.005839686840772629
step: 210, loss: 0.0003140968328807503
step: 220, loss: 0.0003880435542669147
step: 230, loss: 0.045906662940979004
/usr/local/lib/python3.7/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=0.9647577092511014, f1=0.9636163175303197,
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: 0.014832145534455776
step: 10, loss: 0.0023499405942857265
step: 20, loss: 0.0005730282282456756
step: 30, loss: 0.0003760445106308907
step: 40, loss: 0.000288805749733001
step: 50, loss: 0.00044927594717592
step: 60, loss: 0.030798498541116714
step: 70, loss: 0.0006667698035016656
step: 80, loss: 0.00020949533791281283
step: 90, loss: 0.00030408287420868874
step: 100, loss: 0.0002424262638669461
step: 110, loss: 0.00018837297102436423
step: 120, loss: 0.00027770918677560985
step: 130, loss: 0.001486696768552065
step: 140, loss: 0.0002352656447328627
step: 150, loss: 0.03422614559531212
step: 160, loss: 0.0003101770125795156
step: 170, loss: 0.00018353882478550076
step: 180, loss: 0.001127840718254447
step: 190, loss: 0.0005181076703593135
step: 200, loss: 0.0037357984110713005
step: 210, loss: 0.00011175020335940644
step: 220, loss: 0.0002605249173939228
step: 230, loss: 0.0021239807829260826
/usr/local/lib/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=0.9688888888888889, f1=0.9609810479375697,
best_f1=0.9609810479375697
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.008507473394274712
step: 10, loss: 0.00020002250676043332
step: 20, loss: 0.00036712700966745615
step: 30, loss: 0.00016841036267578602
step: 40, loss: 6.606354872928932e-05
step: 50, loss: 0.0001915026514325291
step: 60, loss: 0.000135700567625463
step: 70, loss: 0.0002540086570661515
step: 80, loss: 0.00036863991408608854
step: 90, loss: 0.006389621179550886
step: 100, loss: 0.00010436073353048414
step: 110, loss: 6.675364420516416e-05
step: 120, loss: 0.09054002165794373
step: 130, loss: 0.00016714769299142063
step: 140, loss: 0.03462239354848862
step: 150, loss: 6.25154934823513e-05
step: 160, loss: 0.00017352939175907522
step: 170, loss: 4.959605212206952e-05
step: 180, loss: 0.00012215640163049102
step: 190, loss: 7.247099711094052e-05
step: 200, loss: 8.389534195885062e-05
step: 210, loss: 9.983510244637728e-05
step: 220, loss: 0.0035685489419847727
step: 230, loss: 0.0019472824642434716
/usr/local/lib/python3.7/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=0.9626373626373625, f1=0.9581497797356827,
best_f1=0.9609810479375697
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012271561718080193
step: 10, loss: 9.073880210053176e-05
step: 20, loss: 0.00010598153312457725
step: 30, loss: 0.0001098309803637676
step: 40, loss: 6.82194804539904e-05
step: 50, loss: 0.0007611748296767473
step: 60, loss: 0.015294734388589859
step: 70, loss: 0.0025667273439466953
step: 80, loss: 0.00020876886264886707
step: 90, loss: 0.0009462893358431756
step: 100, loss: 9.535550634609535e-05
step: 110, loss: 0.00010066911636386067
step: 120, loss: 0.0011481497203931212
step: 130, loss: 0.008832765743136406
step: 140, loss: 0.024161875247955322
step: 150, loss: 0.00012315428466536105
```

```
step: 160, loss: 0.0031810400541871786
step: 170, loss: 0.00047740089939907193
step: 180, loss: 0.00010407176159787923
step: 190, loss: 0.00010546597331995144
step: 200, loss: 5.466872244141996e-05
step: 210, loss: 0.00017057110380847007
step: 220, loss: 0.0001126388378906995
step: 230, loss: 0.0002571502991486341
/usr/local/lib/python3.7/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=0.9621380846325166, f1=0.9665178571428571,
best_f1=0.9609810479375697
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010891052079387009
step: 10, loss: 7.957357593113557e-05
step: 20, loss: 9.721935930429026e-05
step: 30, loss: 0.00021943665342405438
step: 40, loss: 0.01961129531264305
step: 50, loss: 0.008305697701871395
step: 60, loss: 0.00045547191984951496
step: 70, loss: 0.00012499264266807586
step: 80, loss: 8.241780597018078e-05
step: 90, loss: 0.01844622567296028
step: 100, loss: 0.00020060574752278626
step: 110, loss: 0.024905789643526077
step: 120, loss: 0.00011606559564825147
step: 130, loss: 5.381866139941849e-05
step: 140, loss: 0.0006831561331637204
step: 150, loss: 5.254490679362789e-05
step: 160, loss: 4.470855128602125e-05
step: 170, loss: 0.00044232106301933527
step: 180, loss: 0.00013629102613776922
step: 190, loss: 9.312136535299942e-05
step: 200, loss: 7.177335646701977e-05
step: 210, loss: 4.889177580480464e-05
step: 220, loss: 5.586783663602546e-05
step: 230, loss: 0.00044596788939088583
/usr/local/lib/python3.7/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=0.9630459126539753, f1=0.967525195968645,
best_f1=0.9609810479375697
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 7.157260552048683e-05
step: 10, loss: 8.909403550205752e-05
step: 20, loss: 4.618719685822725e-05
step: 30, loss: 7.268815534189343e-05
step: 40, loss: 4.8435638746013865e-05
step: 50, loss: 4.7758774599060416e-05
step: 60, loss: 0.00020167882030364126
step: 70, loss: 4.520752918324433e-05
step: 80, loss: 4.5337255869526416e-05
step: 90, loss: 5.9116304328199476e-05
step: 100, loss: 0.00023760694602970034
step: 110, loss: 0.00011689351231325418
step: 120, loss: 7.172991900006309e-05
step: 130, loss: 9.836533718043938e-05
step: 140, loss: 4.4325013732304797e-05
step: 150, loss: 3.552307316567749e-05
step: 160, loss: 0.01698104664683342
step: 170, loss: 6.11983923590742e-05
step: 180, loss: 4.569174052448943e-05
step: 190, loss: 4.16014991060365e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 200, loss: 0.05128265917301178
step: 210, loss: 4.9796792154666036e-05
step: 220, loss: 0.014976835809648037
step: 230, loss: 9.459659486310557e-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=0.9646017699115044, f1=0.9623893805309734,
best_f1=0.9609810479375697
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.860326614812948e-05
step: 10, loss: 7.76608067099005e-05
step: 20, loss: 7.27334336261265e-05
step: 30, loss: 0.00010406381625216454
step: 40, loss: 0.00011238019214943051
step: 50, loss: 6.163004582049325e-05
step: 60, loss: 4.2796800698852167e-05
step: 70, loss: 5.44782196811866e-05
step: 80, loss: 0.0001740532170515507
step: 90, loss: 3.1906020012684166e-05
step: 100, loss: 4.8165682528633624e-05
step: 110, loss: 5.71228374610655e-05
step: 120, loss: 4.8965521273203194e-05
step: 130, loss: 0.00015045094187371433
step: 140, loss: 0.0032290706876665354
```

```
step: 150, loss: 4.6176297473721206e-05
step: 160, loss: 3.900508090737276e-05
step: 170, loss: 0.00010454238508827984
step: 180, loss: 0.150712162256240844
step: 190, loss: 5.529848567675799e-05
step: 200, loss: 6.1052392993588e-05
step: 210, loss: 5.913890709052794e-05
step: 220, loss: 5.14109997311607e-05
step: 230, loss: 3.217056655557826e-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=0.967741935483871, f1=0.967741935483871,
best_f1=0.9609810479375697
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.4606691770022735e-05
step: 10, loss: 2.509669866412878e-05
step: 20, loss: 0.00037381466245278716
step: 30, loss: 6.614817539229989e-05
step: 40, loss: 4.719990101875737e-05
step: 50, loss: 0.00017509258759673685
step: 60, loss: 6.123077037045732e-05
step: 70, loss: 3.803010986302979e-05
step: 80, loss: 0.00032504688715562224
step: 90, loss: 6.933858821867034e-05
step: 100, loss: 0.00023996064555831254
step: 110, loss: 6.485627818619832e-05
step: 120, loss: 4.62389289168641e-05
step: 130, loss: 6.091591058066115e-05
step: 140, loss: 2.982754995173309e-05
step: 150, loss: 5.4006075515644625e-05
step: 160, loss: 3.244612889830023e-05
step: 170, loss: 5.0206639571115375e-05
step: 180, loss: 5.8378824178362265e-05
step: 190, loss: 6.352792115649208e-05
step: 200, loss: 0.00015678134514018893
step: 210, loss: 6.120564648881555e-05
step: 220, loss: 0.00013301869330462068
step: 230, loss: 2.8866586944786832e-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=0.9658965896589659, f1=0.960352422907489,
best_f1=0.9609810479375697
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.461494144285098e-05
step: 10, loss: 6.527360528707504e-05
step: 20, loss: 4.09949898312334e-05
step: 30, loss: 6.104700150899589e-05
step: 40, loss: 6.0989055782556534e-05
step: 50, loss: 5.842745667905547e-05
step: 60, loss: 3.3284202800132334e-05
step: 70, loss: 6.043338726158254e-05
step: 80, loss: 6.23630330665037e-05
step: 90, loss: 3.274378104833886e-05
step: 100, loss: 5.624758341582492e-05
step: 110, loss: 4.921179424854927e-05
step: 120, loss: 4.515185719355941e-05
step: 130, loss: 7.298943819478154e-05
step: 140, loss: 0.010043736547231674
step: 150, loss: 0.0007496593170799315
step: 160, loss: 6.739084346918389e-05
step: 170, loss: 3.017016024386976e-05
step: 180, loss: 4.620952677214518e-05
step: 190, loss: 0.0001074454266927205
step: 200, loss: 4.394101779325865e-05
step: 210, loss: 0.0039021074771881104
step: 220, loss: 0.013012025505304337
step: 230, loss: 5.7081677368842065e-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=0.9658965896589659, f1=0.9623893805309734,
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 [00:08, 277.99it/s]
load_f1 = 0.9670329670329669
real_f1 = 0.963855421686747
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 [00:12, 354.45it/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 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: 0.78961580991745
step: 10, loss: 0.4174054265022278
step: 20, loss: 0.4913294017314911

```
step: 30, loss: 0.41430017352104187
step: 40, loss: 0.3484281897544861
step: 50, loss: 0.21121971309185028
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.2513503134250641
step: 70, loss: 0.22245287895202637
step: 80, loss: 0.2661910355091095
step: 90, loss: 0.25591468811035156
step: 100, loss: 0.22126315534114838
step: 110, loss: 0.19313038885593414
step: 120, loss: 0.04802675545215607
step: 130, loss: 0.03354160115122795
step: 140, loss: 0.3201464116573334
step: 150, loss: 0.03550494462251663
step: 160, loss: 0.23687973618507385
step: 170, loss: 0.35129493474960327
step: 180, loss: 0.10076989978551865
step: 190, loss: 0.07098457962274551
step: 200, loss: 0.10715290904045105
step: 210, loss: 0.10644625127315521
step: 220, loss: 0.2457100749015808
step: 230, loss: 0.1225276067852974
step: 240, loss: 0.10465839505195618
step: 250, loss: 0.11599304527044296
step: 260, loss: 0.04308081790804863
step: 270, loss: 0.023127637803554535
step: 280, loss: 0.13423261046409607
step: 290, loss: 0.07909109443426132
step: 300, loss: 0.1246105208992958
step: 310, loss: 0.09947901219129562
step: 320, loss: 0.09118884056806564
step: 330, loss: 0.11495254188776016
step: 340, loss: 0.1340550035238266
step: 350, loss: 0.1310141682624817
step: 360, loss: 0.06753630936145782
step: 370, loss: 0.10923849046230316
step: 380, loss: 0.2003268152475357
step: 390, loss: 0.04106204956769943
step: 400, loss: 0.036621201783418655
step: 410, loss: 0.05369621515274048
step: 420, loss: 0.026662485674023628
step: 430, loss: 0.10511040687561035
step: 440, loss: 0.10934021323919296
step: 450, loss: 0.08416598290205002
step: 460, loss: 0.018750406801700592
step: 470, loss: 0.15267059206962585
step: 480, loss: 0.23859792947769165
step: 490, loss: 0.07871218025684357
```

```
step: 500, loss: 0.025381801649928093
step: 510, loss: 0.14879655838012695
step: 520, loss: 0.09508290141820908
step: 530, loss: 0.10253879427909851
/usr/local/lib/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=0.9220595181861123, f1=0.9186647860836858,
best_f1=0.9186647860836858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08815119415521622
step: 10, loss: 0.08415036648511887
step: 20, loss: 0.15069472789764404
step: 30, loss: 0.15009362995624542
step: 40, loss: 0.019713718444108963
step: 50, loss: 0.05020301416516304
step: 60, loss: 0.251923143863678
step: 70, loss: 0.10818396508693695
step: 80, loss: 0.019463205710053444
step: 90, loss: 0.04139670357108116
step: 100, loss: 0.15783828496932983
step: 110, loss: 0.03660375624895096
step: 120, loss: 0.07969382405281067
step: 130, loss: 0.03596070036292076
step: 140, loss: 0.04513988271355629
step: 150, loss: 0.049993082880973816
step: 160, loss: 0.010434400290250778
step: 170, loss: 0.09455590695142746
step: 180, loss: 0.06512217968702316
step: 190, loss: 0.16031496226787567
step: 200, loss: 0.07045053690671921
step: 210, loss: 0.06986934691667557
step: 220, loss: 0.1268172711133957
step: 230, loss: 0.052873242646455765
step: 240, loss: 0.1327575445175171
step: 250, loss: 0.07399880141019821
step: 260, loss: 0.010053422302007675
step: 270, loss: 0.07461445033550262
step: 280, loss: 0.3516054153442383
step: 290, loss: 0.04856882989406586
step: 300, loss: 0.1229778304696083
step: 310, loss: 0.05665690451860428
```

```
step: 320, loss: 0.1612665057182312
step: 330, loss: 0.06306253373622894
step: 340, loss: 0.014879006892442703
step: 350, loss: 0.07528968900442123
step: 360, loss: 0.049246661365032196
step: 370, loss: 0.04259604960680008
step: 380, loss: 0.08920049667358398
step: 390, loss: 0.037322722375392914
step: 400, loss: 0.07423800230026245
step: 410, loss: 0.0010179848177358508
step: 420, loss: 0.032247137278318405
step: 430, loss: 0.01644054241478443
step: 440, loss: 0.01357215829193592
step: 450, loss: 0.007066752295941114
step: 460, loss: 0.17622369527816772
step: 470, loss: 0.007351969368755817
step: 480, loss: 0.19446243345737457
step: 490, loss: 0.1450609713792801
step: 500, loss: 0.03131413832306862
step: 510, loss: 0.06132207065820694
step: 520, loss: 0.04185457155108452
step: 530, loss: 0.2509182095527649
/usr/local/lib/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=0.9270588235294117, f1=0.9226441631504922,
best_f1=0.9226441631504922
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009583397768437862
step: 10, loss: 0.01957589201629162
step: 20, loss: 0.13013975322246552
step: 30, loss: 0.18025654554367065
step: 40, loss: 0.011717780493199825
step: 50, loss: 0.014283024705946445
step: 60, loss: 0.029568016529083252
step: 70, loss: 0.008444697596132755
step: 80, loss: 0.0034268724266439676
step: 90, loss: 0.04554666206240654
step: 100, loss: 0.14815661311149597
step: 110, loss: 0.024753950536251068
step: 120, loss: 0.03702748939394951
step: 130, loss: 0.014028563164174557
```

```
step: 140, loss: 0.025959972292218483
step: 150, loss: 0.005014846567064524
step: 160, loss: 0.0036424503196030855
step: 170, loss: 0.053090717643499374
step: 180, loss: 0.0025308094918727875
step: 190, loss: 0.00425513694062829
step: 200, loss: 0.02398955076932907
step: 210, loss: 0.23356938362121582
step: 220, loss: 0.01591157540678978
step: 230, loss: 0.014157116413116455
step: 240, loss: 0.013978616334497929
step: 250, loss: 0.002463612938299775
step: 260, loss: 0.0053888605907559395
step: 270, loss: 0.003102254355326295
step: 280, loss: 0.11643403023481369
step: 290, loss: 0.04884495958685875
step: 300, loss: 0.06212496757507324
step: 310, loss: 0.1566142439842224
step: 320, loss: 0.17157495021820068
step: 330, loss: 0.005277963820844889
step: 340, loss: 0.00418311869725585
step: 350, loss: 0.00885886512696743
step: 360, loss: 0.015221450477838516
step: 370, loss: 0.0208735391497612
step: 380, loss: 0.03950698673725128
step: 390, loss: 0.008254405111074448
step: 400, loss: 0.04007783159613609
step: 410, loss: 0.02681264840066433
step: 420, loss: 0.08486498892307281
step: 430, loss: 0.1306566745042801
step: 440, loss: 0.021592428907752037
step: 450, loss: 0.009009826928377151
step: 460, loss: 0.05830194056034088
step: 470, loss: 0.005567527841776609
step: 480, loss: 0.023993633687496185
step: 490, loss: 0.024288656190037727
step: 500, loss: 0.09120774269104004
step: 510, loss: 0.01113885547965765
step: 520, loss: 0.02030104398727417
step: 530, loss: 0.02683839574456215
/usr/local/lib/python3.7/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=0.9231481481481482, f1=0.919944470152707,
best_f1=0.9226441631504922
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.0046771131455898285
step: 10, loss: 0.006166410632431507
step: 20, loss: 0.06742899119853973
step: 30, loss: 0.010618581436574459
step: 40, loss: 0.007468804717063904
step: 50, loss: 0.03877115994691849
step: 60, loss: 0.002353665651753545
step: 70, loss: 0.0009014091920107603
step: 80, loss: 0.17102932929992676
step: 90, loss: 0.17709262669086456
step: 100, loss: 0.02472306415438652
step: 110, loss: 0.0031535672023892403
step: 120, loss: 0.006232154555618763
step: 130, loss: 0.03281841427087784
step: 140, loss: 0.03351246193051338
step: 150, loss: 0.0037387977354228497
step: 160, loss: 0.009134679101407528
step: 170, loss: 0.045128002762794495
step: 180, loss: 0.019018281251192093
step: 190, loss: 0.002302325563505292
step: 200, loss: 0.0009057965944521129
step: 210, loss: 0.09900958091020584
step: 220, loss: 0.0039284308440983295
step: 230, loss: 0.17136839032173157
step: 240, loss: 0.0035072334576398134
step: 250, loss: 0.027267497032880783
step: 260, loss: 0.11700969934463501
step: 270, loss: 0.01336695160716772
step: 280, loss: 0.012650242075324059
step: 290, loss: 0.01696370542049408
step: 300, loss: 0.002512209117412567
step: 310, loss: 0.001559145632199943
step: 320, loss: 0.18926379084587097
step: 330, loss: 0.030475858598947525
step: 340, loss: 0.025092843919992447
step: 350, loss: 0.0018850177293643355
step: 360, loss: 0.009960032068192959
step: 370, loss: 0.02330918237566948
step: 380, loss: 0.011586425825953484
step: 390, loss: 0.04128705710172653
step: 400, loss: 0.008357676677405834
step: 410, loss: 0.0027645283844321966
step: 420, loss: 0.0334361232817173
step: 430, loss: 0.0075859311036765575
step: 440, loss: 0.04819795489311218
step: 450, loss: 0.0026473680045455694
step: 460, loss: 0.004389037378132343
step: 470, loss: 0.009699279442429543
```

```
step: 480, loss: 0.007264245767146349
step: 490, loss: 0.013731541112065315
step: 500, loss: 0.007177795749157667
step: 510, loss: 0.015943318605422974
step: 520, loss: 0.040843717753887177
step: 530, loss: 0.0019481623312458396
/usr/local/lib/python3.7/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=0.9225746268656716, f1=0.9156626506024096,
best_f1=0.9226441631504922
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006332865450531244
step: 10, loss: 0.0018800162943080068
step: 20, loss: 0.0018641609931364655
step: 30, loss: 0.015656940639019012
step: 40, loss: 0.02500299923121929
step: 50, loss: 0.03523467853665352
step: 60, loss: 0.0007762418827041984
step: 70, loss: 0.0015241751680150628
step: 80, loss: 0.034336432814598083
step: 90, loss: 0.09238380938768387
step: 100, loss: 0.0006936754216440022
step: 110, loss: 0.0007126177079044282
step: 120, loss: 0.003082557115703821
step: 130, loss: 0.0004140788223594427
step: 140, loss: 0.027046628296375275
step: 150, loss: 0.012959626503288746
step: 160, loss: 0.009328977204859257
step: 170, loss: 0.00379372900351882
step: 180, loss: 0.00029516840004362166
step: 190, loss: 0.0004655108496081084
step: 200, loss: 0.00031625168048776686
step: 210, loss: 0.003848845139145851
step: 220, loss: 0.0005508330650627613
step: 230, loss: 0.00042758006020491
step: 240, loss: 0.0062313382513821125
step: 250, loss: 0.020754549652338028
step: 260, loss: 0.0011432903120294213
step: 270, loss: 0.07459966838359833
step: 280, loss: 0.024994470179080963
step: 290, loss: 0.00028161550289951265
step: 300, loss: 0.16563677787780762
step: 310, loss: 0.0018296001944690943
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 320, loss: 0.026934022083878517
```

```
step: 330, loss: 0.0019190119346603751
step: 340, loss: 0.003794587915763259
step: 350, loss: 0.0008393849711865187
step: 360, loss: 0.046758074313402176
step: 370, loss: 0.00291085150092 8402
step: 380, loss: 0.01170192938297987
step: 390, loss: 0.004120661411434412
step: 400, loss: 0.0011595863616093993
step: 410, loss: 0.024935653433203697
step: 420, loss: 0.0028967708349227905
step: 430, loss: 0.02018072083592415
step: 440, loss: 0.002415928989648819
step: 450, loss: 0.015200628899037838
step: 460, loss: 0.0071449363604187965
step: 470, loss: 0.006861105561256409
step: 480, loss: 0.0037507647648453712
step: 490, loss: 0.05443035811185837
step: 500, loss: 0.0016292411601170897
step: 510, loss: 0.006633239332586527
step: 520, loss: 0.029745953157544136
step: 530, loss: 0.01128834206610918
/usr/local/lib/python3.7/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=0.9180633147113593, f1=0.9188687992582291,
best_f1=0.9226441631504922
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003618374466896057
step: 10, loss: 0.009616106748580933
step: 20, loss: 0.000681611942127347
step: 30, loss: 0.0006362751591950655
step: 40, loss: 0.030810635536909103
step: 50, loss: 0.03457128256559372
step: 60, loss: 0.00019575649639591575
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 70, loss: 0.2570161521434784
step: 80, loss: 0.0005665835342369974
step: 90, loss: 0.0006857522530481219
step: 100, loss: 0.0016945298993960023
step: 110, loss: 0.19874140620231628
step: 120, loss: 0.017947033047676086
step: 130, loss: 0.006423383019864559
step: 140, loss: 0.011345569044351578
step: 150, loss: 0.0253542959690094
step: 160, loss: 0.0003830475907307118
step: 170, loss: 0.0006981799961067736
```

```
step: 180, loss: 0.0039058830589056015
step: 190, loss: 0.0017041477840393782
step: 200, loss: 0.0012291503371670842
step: 210, loss: 0.0005598862189799547
step: 220, loss: 0.00039492762880399823
step: 230, loss: 0.0003250178415328264
step: 240, loss: 0.0019721544813364744
step: 250, loss: 0.0014653406105935574
step: 260, loss: 0.0016418294981122017
step: 270, loss: 0.0005146423354744911
step: 280, loss: 0.057668931782245636
step: 290, loss: 0.007670494727790356
step: 300, loss: 0.0005346032558009028
step: 310, loss: 0.06428249925374985
step: 320, loss: 0.0015213062288239598
step: 330, loss: 0.033660512417554855
step: 340, loss: 0.059377145022153854
step: 350, loss: 0.1100664883852005
step: 360, loss: 0.0023098294623196125
step: 370, loss: 0.006367037538439035
step: 380, loss: 0.0013320555444806814
step: 390, loss: 0.0037411709781736135
step: 400, loss: 0.0005301603814586997
step: 410, loss: 0.0070157176814973354
step: 420, loss: 0.055441223084926605
step: 430, loss: 0.024119216948747635
step: 440, loss: 0.07040081918239594
step: 450, loss: 0.0175349880098896
step: 460, loss: 0.008355592377483845
step: 470, loss: 0.06585802882909775
step: 480, loss: 0.019548170268535614
step: 490, loss: 0.0005463496199809015
step: 500, loss: 0.0005002242978662252
step: 510, loss: 0.0003025658952537924
step: 520, loss: 0.0006456485134549439
step: 530, loss: 0.001802901504561305
/usr/local/lib/python3.7/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=0.9212635549269212, f1=0.9256820319849483,
best_f1=0.9226441631504922
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018744097324088216
step: 10, loss: 0.017104627564549446
step: 20, loss: 0.006000048480927944
step: 30, loss: 0.02962174080312252
```

```
step: 40, loss: 0.0010627121664583683
step: 50, loss: 0.0008624998736195266
step: 60, loss: 0.00046309910248965025
step: 70, loss: 0.0011661805910989642
step: 80, loss: 0.002570117823779583
step: 90, loss: 0.0001565551065141335
step: 100, loss: 0.0022134489845484495
step: 110, loss: 0.005514821968972683
step: 120, loss: 0.00016342122398782521
step: 130, loss: 0.0015322132967412472
step: 140, loss: 0.0003905505873262882
step: 150, loss: 0.00015582430933136493
step: 160, loss: 0.0014714201679453254
step: 170, loss: 0.00304427114315331
step: 180, loss: 0.015120877884328365
step: 190, loss: 0.028033079579472542
step: 200, loss: 0.0031708842143416405
step: 210, loss: 0.0011350185377523303
step: 220, loss: 0.0003367544268257916
step: 230, loss: 0.0003383637813385576
step: 240, loss: 0.0029958803206682205
step: 250, loss: 0.005080558825284243
step: 260, loss: 0.000764877418987453
step: 270, loss: 0.1128271147608757
step: 280, loss: 0.001479415688663721
step: 290, loss: 0.000739060458727181
step: 300, loss: 0.0011711729457601905
step: 310, loss: 0.034294094890356064
step: 320, loss: 0.013964518904685974
step: 330, loss: 0.0004499058413784951
step: 340, loss: 0.31401780247688293
step: 350, loss: 0.07306022942066193
step: 360, loss: 0.0020301868207752705
step: 370, loss: 0.002392926486209035
step: 380, loss: 0.0021214636508375406
step: 390, loss: 0.0005315249436534941
step: 400, loss: 0.003684005467221141
step: 410, loss: 0.004279326181858778
step: 420, loss: 0.0042830598540604115
step: 430, loss: 0.001483267406001687
step: 440, loss: 0.0003613524604588747
step: 450, loss: 0.0038786993827670813
step: 460, loss: 0.00241479161195457
step: 470, loss: 0.16348743438720703
step: 480, loss: 0.0009184449445456266
step: 490, loss: 0.0013375479029491544
step: 500, loss: 0.001445053261704743
step: 510, loss: 0.002843126654624939
```

```
step: 520, loss: 0.1100078597664833
step: 530, loss: 0.023776637390255928
/usr/local/lib/python3.7/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=0.9237209302325581, f1=0.9227209625173529,
best_f1=0.9226441631504922
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00038254508399404585
step: 10, loss: 0.004742704797536135
step: 20, loss: 0.013916981406509876
step: 30, loss: 0.0025647038128226995
step: 40, loss: 0.005107062868773937
step: 50, loss: 0.0011970868799835443
step: 60, loss: 0.003664359450340271
step: 70, loss: 0.133827805519104
step: 80, loss: 0.0004294113605283201
step: 90, loss: 0.00042130169458687305
step: 100, loss: 0.0008055659127421677
step: 110, loss: 0.0011562107829377055
step: 120, loss: 0.013346867635846138
step: 130, loss: 0.00021608623501379043
step: 140, loss: 0.00013609914458356798
step: 150, loss: 0.0007211975171230733
step: 160, loss: 0.00231330469250679
step: 170, loss: 0.0005061249248683453
step: 180, loss: 0.0004980401718057692
step: 190, loss: 0.004336107987910509
step: 200, loss: 0.0038567082956433296
step: 210, loss: 0.009336366318166256
step: 220, loss: 0.009402497671544552
step: 230, loss: 0.08124230802059174
step: 240, loss: 0.0006873546517454088
step: 250, loss: 0.0006008355412632227
step: 260, loss: 0.0015138498274609447
step: 270, loss: 0.00010747499618446454
step: 280, loss: 0.0007434425642713904
step: 290, loss: 0.0063987611792981625
step: 300, loss: 0.00036329554859548807
step: 310, loss: 0.029977694153785706
step: 320, loss: 0.0022547102998942137
step: 330, loss: 0.11319062858819962
step: 340, loss: 0.00026622533914633095
step: 350, loss: 0.005453978665173054
step: 360, loss: 0.0004979336517862976
step: 370, loss: 0.00022835787967778742
```

```
step: 380, loss: 0.0003669953439384699
step: 390, loss: 0.05232849717140198
step: 400, loss: 0.005978364497423172
step: 410, loss: 0.024261312559247017
step: 420, loss: 0.0002972203947138041
step: 430, loss: 0.00039214035496115685
step: 440, loss: 0.008639290928840637
step: 450, loss: 0.0004424079670570791
step: 460, loss: 0.0028161415830254555
step: 470, loss: 0.0004216067318338901
step: 480, loss: 0.0007044707890599966
step: 490, loss: 0.00017652631504461117
step: 500, loss: 0.0007130582234822214
step: 510, loss: 0.0015204788651317358
step: 520, loss: 0.00036697665927931666
step: 530, loss: 0.00024146946088876575
/usr/local/lib/python3.7/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=0.9242990654205607, f1=0.9223616922361693,
best_f1=0.9226441631504922
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013783469330519438
step: 10, loss: 0.09717255085706711
step: 20, loss: 0.0004850584373343736
step: 30, loss: 0.0006872583762742579
step: 40, loss: 0.009330407716333866
step: 50, loss: 0.0009286439744755626
step: 60, loss: 0.0006524131167680025
step: 70, loss: 0.06320040673017502
step: 80, loss: 0.0005733669968321919
step: 90, loss: 0.03432145714759827
step: 100, loss: 0.00042340761865489185
step: 110, loss: 0.003990076016634703
step: 120, loss: 9.112631232710555e-05
step: 130, loss: 0.0009524478809908032
step: 140, loss: 0.058910444378852844
step: 150, loss: 0.0001689391501713544
step: 160, loss: 0.0002434096677461639
step: 170, loss: 0.0001649269397603348
step: 180, loss: 0.00012525010970421135
step: 190, loss: 0.0008580504218116403
step: 200, loss: 0.0012570885010063648
step: 210, loss: 0.00011551391071407124
step: 220, loss: 0.00025120400823652744
step: 230, loss: 7.64008509577252e-05
```

```
step: 240, loss: 0.0003537914017215371
step: 250, loss: 0.0002520094276405871
step: 260, loss: 0.00014065652794670314
step: 270, loss: 0.0036707029212266207
step: 280, loss: 0.0001303235476370901
step: 290, loss: 0.0005496650701388717
step: 300, loss: 0.00014867426943965256
step: 310, loss: 6.146441592136398e-05
step: 320, loss: 0.0031624794793767023
step: 330, loss: 0.0007794919074513018
step: 340, loss: 0.0002782887313514948
step: 350, loss: 8.138268458424136e-05
step: 360, loss: 0.03622376546263695
step: 370, loss: 0.0002942031132986317
step: 380, loss: 8.039540989557281e-05
step: 390, loss: 0.00018459193233866245
step: 400, loss: 0.002424542559310794
step: 410, loss: 0.00028381627635098994
step: 420, loss: 0.0005986217875033617
step: 430, loss: 7.010501576587558e-05
step: 440, loss: 0.042120400816202164
step: 450, loss: 0.00010593080514809117
step: 460, loss: 0.00015373778296634555
step: 470, loss: 0.00010715817916207016
step: 480, loss: 0.002101604361087084
step: 490, loss: 0.00014583593292627484
step: 500, loss: 0.00025901314802467823
step: 510, loss: 0.0009867037879303098
step: 520, loss: 9.609747212380171e-05
step: 530, loss: 8.209012594306841e-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=0.9214758751182593,
best_f1=0.9226441631504922
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00026285176863893867
step: 10, loss: 0.00020798921468667686
step: 20, loss: 0.00019318723934702575
step: 30, loss: 7.381034083664417e-05
step: 40, loss: 0.0008176541887223721
step: 50, loss: 7.317098788917065e-05
step: 60, loss: 9.183173824567348e-05
step: 70, loss: 0.00020876289636362344
step: 80, loss: 0.0010130748885122478
step: 90, loss: 0.00018473717500455678
```

```
step: 100, loss: 0.00012806206359528005
step: 110, loss: 0.011171232908964157
step: 120, loss: 0.00011226595233893022
step: 130, loss: 0.00015367595187854022
step: 140, loss: 0.0005204788176342845
step: 150, loss: 0.0002890351170208305
step: 160, loss: 0.00015664342208765447
step: 170, loss: 0.0007279720739461482
step: 180, loss: 0.00021789160382468253
step: 190, loss: 0.0010885841911658645
step: 200, loss: 0.00042682865750975907
step: 210, loss: 8.52470111567527e-05
step: 220, loss: 6.175833550514653e-05
step: 230, loss: 0.00011984312004642561
step: 240, loss: 6.06358953518793e-05
step: 250, loss: 0.00027226778911426663
step: 260, loss: 0.020582664757966995
step: 270, loss: 0.07732041925191879
step: 280, loss: 7.779137376928702e-05
step: 290, loss: 6.715965719195083e-05
step: 300, loss: 0.0003840526915155351
step: 310, loss: 7.584834384033456e-05
step: 320, loss: 5.2256116759963334e-05
step: 330, loss: 0.00025210948660969734
step: 340, loss: 8.20812419988215e-05
step: 350, loss: 0.000132224740809761
step: 360, loss: 9.644714009482414e-05
step: 370, loss: 0.00023599533597007394
step: 380, loss: 0.00011773137521231547
step: 390, loss: 0.0002815878251567483
step: 400, loss: 9.506393689662218e-05
step: 410, loss: 0.0002575321705080569
step: 420, loss: 0.0002163832978112623
step: 430, loss: 0.00010983360698446631
step: 440, loss: 0.00012884444473175332
step: 450, loss: 0.0013508135452866554
step: 460, loss: 0.03213245049118996
step: 470, loss: 8.004142728168517e-05
step: 480, loss: 0.002299876417964697
step: 490, loss: 0.3104783594608307
step: 500, loss: 0.035769566893577576
step: 510, loss: 0.016622869297862053
step: 520, loss: 0.07274418324232101
step: 530, loss: 0.0001526513951830566
/usr/local/lib/python3.7/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=0.9210275927687916, f1=0.922279792746114,
```

```
best_f1=0.9226441631504922
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003237581695429981
step: 10, loss: 0.0013254276709631085
step: 20, loss: 0.0001282974990317598
step: 30, loss: 0.001265497412532568
step: 40, loss: 0.000224292220082134
step: 50, loss: 0.020653609186410904
step: 60, loss: 0.00026365145458839834
step: 70, loss: 0.0003990136901848018
step: 80, loss: 0.004329978954046965
step: 90, loss: 0.0101699847728014
step: 100, loss: 0.0005802419036626816
step: 110, loss: 0.00029741544858552516
step: 120, loss: 0.000483192503452301
step: 130, loss: 0.00018012626969721168
step: 140, loss: 0.00014592551451642066
step: 150, loss: 7.637857925146818e-05
step: 160, loss: 0.0005997923435643315
step: 170, loss: 0.0006331123295240104
step: 180, loss: 0.0001163749911938794
step: 190, loss: 0.00033621222246438265
step: 200, loss: 0.0002200679009547457
step: 210, loss: 0.00014597379777114838
step: 220, loss: 0.0006440077559091151
step: 230, loss: 0.001209234818816185
step: 240, loss: 7.38351300242357e-05
step: 250, loss: 0.0003340683178976178
step: 260, loss: 9.436998516321182e-05
step: 270, loss: 0.005773558747023344
step: 280, loss: 0.00017313404532615095
step: 290, loss: 0.0002990917128045112
step: 300, loss: 0.0010299109853804111
step: 310, loss: 0.0002724690712057054
step: 320, loss: 0.0016389822121709585
step: 330, loss: 0.00012028494529658929
step: 340, loss: 8.260799950221553e-05
step: 350, loss: 0.00010247422324027866
step: 360, loss: 0.0001325156190432608
step: 370, loss: 0.00015085494669619948
step: 380, loss: 0.00020931308972649276
step: 390, loss: 0.0007072677835822105
step: 400, loss: 0.0003439475258346647
step: 410, loss: 7.968365389388055e-05
step: 420, loss: 0.0001848226966103539
step: 430, loss: 0.0003160572377964854
```

```
step: 440, loss: 7.934436871437356e-05
step: 450, loss: 0.00013714110536966473
step: 460, loss: 0.0002178701979573816
step: 470, loss: 0.0018945670453831553
step: 480, loss: 0.00018699435167945921
step: 490, loss: 0.002334253629669547
step: 500, loss: 0.00019133754540234804
step: 510, loss: 9.408679761691019e-05
step: 520, loss: 9.280556696467102e-05
step: 530, loss: 0.0003440168802626431
/usr/local/lib/python3.7/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=0.9250936329588015, f1=0.9221501390176089,
best_f1=0.9226441631504922
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006722760736010969
step: 10, loss: 5.0508529966464266e-05
step: 20, loss: 3.2535361242480576e-05
step: 30, loss: 9.859087003860623e-05
step: 40, loss: 0.00015194567095022649
step: 50, loss: 0.0030165507923811674
step: 60, loss: 0.00019035184232052416
step: 70, loss: 0.0033325066324323416
step: 80, loss: 5.799972859676927e-05
step: 90, loss: 5.74682853766717e-05
step: 100, loss: 0.0002206295175710693
step: 110, loss: 0.0015787709271535277
step: 120, loss: 0.00014335148443933576
step: 130, loss: 0.00010633825149852782
step: 140, loss: 6.569752440555021e-05
step: 150, loss: 0.0003580380289349705
step: 160, loss: 0.00020493980264291167
step: 170, loss: 5.236572178546339e-05
step: 180, loss: 0.00016947522817645222
step: 190, loss: 8.887080184649676e-05
step: 200, loss: 7.51040643081069e-05
step: 210, loss: 0.00022393863764591515
step: 220, loss: 3.969867975683883e-05
step: 230, loss: 0.0001513360912213102
step: 240, loss: 0.0001374351850245148
step: 250, loss: 0.0001245441526407376
step: 260, loss: 0.0006149148102849722
step: 270, loss: 5.7383731473237276e-05
step: 280, loss: 0.00014420674415305257
step: 290, loss: 0.00012146566587034613
```

```
step: 300, loss: 0.0001273404195671901
step: 310, loss: 0.00015624077059328556
step: 320, loss: 5.795125616714358e-05
step: 330, loss: 4.112413807888515e-05
step: 340, loss: 3.3854412322398275e-05
step: 350, loss: 0.041469279676675797
step: 360, loss: 0.00022090481070335954
step: 370, loss: 0.00011758141772588715
step: 380, loss: 7.38723756512627e-05
step: 390, loss: 3.301991455373354e-05
step: 400, loss: 3.304210986243561e-05
step: 410, loss: 0.0004327007045503706
step: 420, loss: 4.421141420607455e-05
step: 430, loss: 0.0008512631757184863
step: 440, loss: 0.0007932159351184964
step: 450, loss: 0.00036056817043572664
step: 460, loss: 4.733836613013409e-05
step: 470, loss: 3.4591605071909726e-05
step: 480, loss: 0.00016918015899136662
step: 490, loss: 5.588128260569647e-05
step: 500, loss: 8.623366738902405e-05
step: 510, loss: 5.704432260245085e-05
step: 520, loss: 0.00024168912204913795
step: 530, loss: 5.4284399084281176e-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=0.9253871421867668, f1=0.9231485794131347,
best_f1=0.9226441631504922
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006250685546547174
step: 10, loss: 2.5878738597384654e-05
step: 20, loss: 0.00024472756194882095
step: 30, loss: 0.00022226515284273773
step: 40, loss: 3.0475814128294587e-05
step: 50, loss: 4.344881381257437e-05
step: 60, loss: 0.0005180949810892344
step: 70, loss: 3.881541488226503e-05
step: 80, loss: 2.9041359084658325e-05
step: 90, loss: 8.933465869631618e-05
step: 100, loss: 2.1133195332367904e-05
step: 110, loss: 5.032621629652567e-05
step: 120, loss: 5.8825808082474396e-05
step: 130, loss: 7.382366311503574e-05
step: 140, loss: 0.00010077707702293992
step: 150, loss: 3.1041799957165495e-05
```

```
step: 160, loss: 0.0005787329864688218
step: 170, loss: 7.713065133430064e-05
step: 180, loss: 3.0132607207633555e-05
step: 190, loss: 3.685301635414362e-05
step: 200, loss: 2.200119342887774e-05
step: 210, loss: 3.046407800866291e-05
step: 220, loss: 7.117405766621232e-05
step: 230, loss: 0.00025816282141022384
step: 240, loss: 4.187343074590899e-05
step: 250, loss: 0.04378648102283478
step: 260, loss: 9.367776510771364e-05
step: 270, loss: 5.116290776641108e-05
step: 280, loss: 0.0008975251112133265
step: 290, loss: 3.730220487341285e-05
step: 300, loss: 0.0002565115864854306
step: 310, loss: 5.937572495895438e-05
step: 320, loss: 4.91429163957946e-05
step: 330, loss: 0.00024953679530881345
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 340, loss: 2.662027873157058e-05
step: 350, loss: 0.005386717151850462
step: 360, loss: 4.547238495433703e-05
step: 370, loss: 2.9800117772538215e-05
step: 380, loss: 4.4252075895201415e-05
step: 390, loss: 7.899617776274681e-05
step: 400, loss: 2.692511770874262e-05
step: 410, loss: 2.925375520135276e-05
step: 420, loss: 2.8002024919260293e-05
step: 430, loss: 0.004594902973622084
step: 440, loss: 3.8971673347987235e-05
step: 450, loss: 4.297250779927708e-05
step: 460, loss: 2.581178705440834e-05
step: 470, loss: 0.00015929726941976696
step: 480, loss: 3.327973900013603e-05
step: 490, loss: 0.0001011893061548471
step: 500, loss: 4.492872903938405e-05
step: 510, loss: 3.13847376673948e-05
step: 520, loss: 0.00017093429050873965
step: 530, loss: 4.338189683039673e-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 13: dev_f1=0.9273066169617894, f1=0.9239280774550485,
best_f1=0.9239280774550485
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.554364179261029e-05
step: 10, loss: 0.00010516018664930016
step: 20, loss: 2.1751586245954968e-05
step: 30, loss: 4.408675886224955e-05
step: 40, loss: 5.511871859198436e-05
step: 50, loss: 0.0008265821961686015
step: 60, loss: 4.0331215132027864e-05
step: 70, loss: 3.892096356139518e-05
step: 80, loss: 3.9944225136423483e-05
step: 90, loss: 4.130430897930637e-05
step: 100, loss: 0.04759576916694641
step: 110, loss: 2.7104559194413014e-05
step: 120, loss: 0.012369070202112198
step: 130, loss: 2.3811464416212402e-05
step: 140, loss: 0.002138382289558649
step: 150, loss: 7.169567834353074e-05
step: 160, loss: 0.000434037676313892
step: 170, loss: 0.07671765983104706
step: 180, loss: 0.00036086354521103203
step: 190, loss: 0.00016068473632913083
step: 200, loss: 4.476389949559234e-05
step: 210, loss: 0.01485525630414486
step: 220, loss: 2.9056229323032312e-05
step: 230, loss: 2.6843797968467698e-05
step: 240, loss: 3.5940014640800655e-05
step: 250, loss: 3.2401603675680235e-05
step: 260, loss: 2.2999120119493455e-05
step: 270, loss: 4.7404420911334455e-05
step: 280, loss: 2.7904921807930805e-05
step: 290, loss: 3.838170232484117e-05
step: 300, loss: 4.8582867748336866e-05
step: 310, loss: 4.271028228686191e-05
step: 320, loss: 0.0012158815516158938
step: 330, loss: 3.483988257357851e-05
step: 340, loss: 3.717291110660881e-05
step: 350, loss: 4.949043795932084e-05
step: 360, loss: 2.5685243599582464e-05
step: 370, loss: 2.973023038066458e-05
step: 380, loss: 5.037014852860011e-05
step: 390, loss: 2.9458431527018547e-05
step: 400, loss: 6.388915062416345e-05
step: 410, loss: 0.00017944256251221221
step: 420, loss: 2.380416299274657e-05
step: 430, loss: 3.4982360375579447e-05
step: 440, loss: 5.61944289074745e-05
```

```
step: 450, loss: 3.4501805203035474e-05
step: 460, loss: 0.00014000666851643473
step: 470, loss: 4.3364874727558345e-05
step: 480, loss: 4.3275009375065565e-05
step: 490, loss: 3.0307552151498385e-05
step: 500, loss: 9.468153439229354e-05
step: 510, loss: 3.9347793062916026e-05
step: 520, loss: 5.374056490836665e-05
step: 530, loss: 3.707043651957065e-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 14: dev_f1=0.9288040949278734, f1=0.9236430542778289,
best_f1=0.9236430542778289
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.969402551883832e-05
step: 10, loss: 4.3403164454502985e-05
step: 20, loss: 4.418440948938951e-05
step: 30, loss: 3.984336944995448e-05
step: 40, loss: 1.9594308469095267e-05
step: 50, loss: 2.7952684831689112e-05
step: 60, loss: 4.718289346783422e-05
step: 70, loss: 2.8571059374371544e-05
step: 80, loss: 3.025540354428813e-05
step: 90, loss: 2.2954542146180756e-05
step: 100, loss: 2.4373646738240495e-05
step: 110, loss: 0.0014285990037024021
step: 120, loss: 2.4284590836032294e-05
step: 130, loss: 2.4251006834674627e-05
step: 140, loss: 2.102142207149882e-05
step: 150, loss: 4.48797109129373e-05
step: 160, loss: 4.471638021641411e-05
step: 170, loss: 2.9112105039530434e-05
step: 180, loss: 2.0987758034607396e-05
step: 190, loss: 7.884797378210351e-05
step: 200, loss: 4.420469849719666e-05
step: 210, loss: 4.599398744176142e-05
step: 220, loss: 1.7419233699911274e-05
step: 230, loss: 0.00042431586189195514
step: 240, loss: 3.677798304124735e-05
step: 250, loss: 7.398879097308964e-05
step: 260, loss: 3.456140620983206e-05
```

```
step: 270, loss: 8.091023482847959e-05
step: 280, loss: 4.3845582695212215e-05
step: 290, loss: 3.756796650122851e-05
step: 300, loss: 0.00010770049266284332
step: 310, loss: 3.788204048760235e-05
step: 320, loss: 0.00020901132666040212
step: 330, loss: 6.159784970805049e-05
step: 340, loss: 0.001052030362188816
step: 350, loss: 4.791283936356194e-05
step: 360, loss: 4.458452531252988e-05
step: 370, loss: 1.9412163965171203e-05
step: 380, loss: 4.0293962229043245e-05
step: 390, loss: 0.0012340389657765627
step: 400, loss: 0.0019840977620333433
step: 410, loss: 0.0001445574453100562
step: 420, loss: 4.052509757457301e-05
step: 430, loss: 2.5796862246352248e-05
step: 440, loss: 0.00499862851575017
step: 450, loss: 9.166399104287848e-05
step: 460, loss: 1.6696521925041452e-05
step: 470, loss: 1.982561116165016e-05
step: 480, loss: 2.772858897515107e-05
step: 490, loss: 0.00526882028758526
step: 500, loss: 3.609613122534938e-05
step: 510, loss: 6.637726619374007e-05
step: 520, loss: 4.7001081838971004e-05
step: 530, loss: 2.857564868463669e-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=0.9283054003724395, f1=0.9235023041474655,
best_f1=0.9236430542778289
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.")
5742it [00:17, 337.08it/s]
load_f1 = 0.9214218896164639
real_f1 = 0.9240093240093239
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 [00:12, 365.83it/s]
```

4   DITTO

4.1   DITTO STRUCTURED

4.1.1   Beer - Running the matcher

```python
# 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 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: 0.8642251491546631
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=0.43478260869565216, f1=0.2608695652173913,
best_f1=0.2608695652173913
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3628312945365906
/usr/local/lib/python3.7/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=0.3333333333333333, f1=0.3846153846153846,
best_f1=0.2608695652173913
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.34026798605918884
/usr/local/lib/python3.7/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=0.3111111111111111, f1=0.29885057471264365,
best_f1=0.2608695652173913
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3566122055053711
/usr/local/lib/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=0.5405405405405405, f1=0.27272727272727276,
best_f1=0.27272727272727276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30580267310142517
/usr/local/lib/python3.7/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=0.5294117647058824, f1=0.3902439024390244,
best_f1=0.27272727272727276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3048407733440399
/usr/local/lib/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=0.5555555555555556, f1=0.42553191489361697,
best_f1=0.42553191489361697
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21200688183307648
/usr/local/lib/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=0.6086956521739131, f1=0.32, best_f1=0.32
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3070550560951233
/usr/local/lib/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=0.6666666666666667, f1=0.4878048780487805,
best_f1=0.4878048780487805
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1371927261352539
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.")
epoch 9: dev_f1=0.6206896551724138, f1=0.5000000000000001,
best_f1=0.4878048780487805
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3674279749393463
/usr/local/lib/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=0.7027027027027025, f1=0.5263157894736842,
best_f1=0.5263157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.229985773563385
/usr/local/lib/python3.7/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=0.6206896551724138, f1=0.48484848484848486,
best_f1=0.5263157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1171126440167427
/usr/local/lib/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=0.7142857142857143, f1=0.47058823529411764,
best_f1=0.47058823529411764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06502936780452728
/usr/local/lib/python3.7/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=0.689655172413793, f1=0.5555555555555556,
best_f1=0.47058823529411764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1273263692855835
/usr/local/lib/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 14: dev_f1=0.7333333333333334, f1=0.5555555555555556,
best_f1=0.5555555555555556
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12028998881578445
/usr/local/lib/python3.7/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=0.7333333333333334, f1=0.5555555555555556,
best_f1=0.5555555555555556
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, 129208.42it/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 = 0.65
real_f1 = 0.6190476190476191
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 [00:12, 358.68it/s]
```

4.1.2 DBLP-ACM - Running the matcher

```python
# 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 \
```

360

```
--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: 0.8014160990715027
step: 10, loss: 0.4671550691127777
step: 20, loss: 0.5835999846458435
step: 30, loss: 0.47365593910217285
step: 40, loss: 0.32909727096557617
```

```
step: 50, loss: 0.13736671209335327
step: 60, loss: 0.15688486397266388
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 70, loss: 0.17084617912769318
step: 80, loss: 0.1744384616613388
step: 90, loss: 0.029224518686532974
step: 100, loss: 0.11246838420629501
step: 110, loss: 0.018786529079079628
step: 120, loss: 0.023186206817626953
step: 130, loss: 0.03203357383608818
step: 140, loss: 0.0869114026427269
step: 150, loss: 0.04214215278625488
step: 160, loss: 0.25485554337501526
step: 170, loss: 0.008733264170587063
step: 180, loss: 0.01040993444621563
step: 190, loss: 0.02385893277823925
step: 200, loss: 0.007578188553452492
step: 210, loss: 0.021241333335638046
step: 220, loss: 0.02896767109632492
step: 230, loss: 0.06959240883588791
/usr/local/lib/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=0.9755555555555556, f1=0.9753914988814317,
best_f1=0.9753914988814317
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018611133985221386
step: 10, loss: 0.0011974603403359652
step: 20, loss: 0.003783273044973612
step: 30, loss: 0.017753994092345238
step: 40, loss: 0.002621033461764455
step: 50, loss: 0.0779181495308876
step: 60, loss: 0.010133251547813416
step: 70, loss: 0.07125749439001083
step: 80, loss: 0.004890927579253912
step: 90, loss: 0.022214045748114586
step: 100, loss: 0.06458917260169983
step: 110, loss: 0.004420396406203508
step: 120, loss: 0.0018130084499716759
step: 130, loss: 0.0037578539922833443
step: 140, loss: 0.025449702516198158
step: 150, loss: 0.04571662098169327
```

```
step: 160, loss: 0.00563557306304574
step: 170, loss: 0.06594028323888779
step: 180, loss: 0.0017886788118630648
step: 190, loss: 0.19266954064369202
step: 200, loss: 0.0555744506418705
step: 210, loss: 0.0440034493803978
step: 220, loss: 0.002163850935176015
step: 230, loss: 0.11279454827308655
/usr/local/lib/python3.7/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=0.9727891156462585, f1=0.9703196347031963,
best_f1=0.9753914988814317
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05380111560225487
step: 10, loss: 0.013071835041046143
step: 20, loss: 0.010849224403500557
step: 30, loss: 0.019599806517362595
step: 40, loss: 0.024661598727107048
step: 50, loss: 0.001298832125030458
step: 60, loss: 0.0629279613494873
step: 70, loss: 0.0008668597438372672
step: 80, loss: 0.002254005055874586
step: 90, loss: 0.007051941938698292
step: 100, loss: 0.0016828600782901049
step: 110, loss: 0.00350264017470181
step: 120, loss: 0.012991119176149368
step: 130, loss: 0.003810547525063157
step: 140, loss: 0.009431286714971066
step: 150, loss: 0.0017625405453145504
step: 160, loss: 0.006260917987674475
step: 170, loss: 0.009446915239095688
step: 180, loss: 0.0013196407817304134
step: 190, loss: 0.015112190507352352
step: 200, loss: 0.002758781658485532
step: 210, loss: 0.02338814176619053
step: 220, loss: 0.040214911103248596
step: 230, loss: 0.06610088795423508
/usr/local/lib/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=0.9855072463768116, f1=0.9787709497206705,
```

best_f1=0.9787709497206705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00836907234042883
step: 10, loss: 0.00482561718672514
step: 20, loss: 0.04233841225504875
step: 30, loss: 0.0010378701845183969
step: 40, loss: 0.013634826987981796
step: 50, loss: 0.005112132523208857
step: 60, loss: 0.00070750457234680065
step: 70, loss: 0.1801706850528717
step: 80, loss: 0.027029529213905334
step: 90, loss: 0.04361213743686676
step: 100, loss: 0.0009702310198917985
step: 110, loss: 0.0448891781270504
step: 120, loss: 0.0017993072979152203
step: 130, loss: 0.005344079341739416
step: 140, loss: 0.0014499437529593706
step: 150, loss: 0.003289218293502927
step: 160, loss: 0.0002288538235006854
step: 170, loss: 0.008588526397943497
step: 180, loss: 0.029635505750775337
step: 190, loss: 0.017226463183760643
step: 200, loss: 0.0008754654554650187
step: 210, loss: 0.00039489931077696383
step: 220, loss: 0.01513802632689476
step: 230, loss: 0.0002897457161452621
/usr/local/lib/python3.7/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=0.9776286353467561, f1=0.9808342728297633,
best_f1=0.9787709497206705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003056522982660681
step: 10, loss: 0.000588955357670784
step: 20, loss: 0.00020685233175754547
step: 30, loss: 0.00024201892665587366
step: 40, loss: 0.0005718299653381109
step: 50, loss: 0.0001999780215555802
step: 60, loss: 0.0018829393666237593
step: 70, loss: 0.07435651868581772
step: 80, loss: 0.01732424460351467
step: 90, loss: 0.037219755351543427
step: 100, loss: 0.03513072803616524
step: 110, loss: 0.11922655254602432

```
step: 120, loss: 0.03225510194897652
step: 130, loss: 0.044226039201021194
step: 140, loss: 0.0002478045062161982
step: 150, loss: 0.00030494111706502736
step: 160, loss: 0.024824626743793488
step: 170, loss: 0.026689445599913597
step: 180, loss: 0.0005117456894367933
step: 190, loss: 0.001876283437013626
step: 200, loss: 0.0016735787503421307
step: 210, loss: 0.0012529956875368953
step: 220, loss: 0.00018159733735956252
step: 230, loss: 0.00029629477648995817
/usr/local/lib/python3.7/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=0.9818594104308391, f1=0.9725400457665903,
best_f1=0.9787709497206705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003873155452311039
step: 10, loss: 0.0004489455313887447
step: 20, loss: 0.023416852578520775
step: 30, loss: 0.013186758384108543
step: 40, loss: 0.0003909761144313961
step: 50, loss: 0.02110920287668705
step: 60, loss: 0.016576189547777176
step: 70, loss: 0.00034963968209922314
step: 80, loss: 0.007219030987471342
step: 90, loss: 0.00019366125343367457
step: 100, loss: 0.00023136267554946244
step: 110, loss: 0.14431636035442352
step: 120, loss: 0.000376890558982268
step: 130, loss: 0.00153874431271106
step: 140, loss: 0.0009335512877441943
step: 150, loss: 0.0008242076728492975
step: 160, loss: 0.003783243475481868
step: 170, loss: 0.0016365647315979004
step: 180, loss: 0.0024037742987275124
step: 190, loss: 0.0019225013675168157
step: 200, loss: 0.0051176464185118675
step: 210, loss: 0.003912059124559164
step: 220, loss: 9.519890591036528e-05
step: 230, loss: 0.001510045723989606
/usr/local/lib/python3.7/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=0.9841628959276018, f1=0.9775280898876404,
```

```
best_f1=0.9787709497206705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.047423675656318665
step: 10, loss: 0.00019041680207010359
step: 20, loss: 7.831087714293972e-05
step: 30, loss: 0.0023964024148881435
step: 40, loss: 0.0008149040513671935
step: 50, loss: 7.284952880581841e-05
step: 60, loss: 0.0202962867915653034
step: 70, loss: 0.0004317510756663978
step: 80, loss: 0.00025449550594203174
step: 90, loss: 0.00917512271553278
step: 100, loss: 0.0017024630215018988
step: 110, loss: 0.0028877367731183767
step: 120, loss: 0.00025931629352271557
step: 130, loss: 0.00017653887334745377
step: 140, loss: 0.0003265165432821959
step: 150, loss: 0.0018204678781330585
step: 160, loss: 0.0013127276906743646
step: 170, loss: 0.006177361123263836
step: 180, loss: 0.08060581982135773
step: 190, loss: 0.016657672822475433
step: 200, loss: 0.0003342133422847837
step: 210, loss: 0.00012922794849146158
step: 220, loss: 0.00029653406818397343
step: 230, loss: 0.026673205196857452
/usr/local/lib/python3.7/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=0.9829351535836178, f1=0.9773755656108598,
best_f1=0.9787709497206705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00946663785725832
step: 10, loss: 0.11240183562040329
step: 20, loss: 0.005102846771478653
step: 30, loss: 0.06203622743487358
step: 40, loss: 0.00016022531781345606
step: 50, loss: 0.001484235399402678
step: 60, loss: 0.00015526977949775755
step: 70, loss: 0.00016739178681746125
step: 80, loss: 0.00022259214892983437
step: 90, loss: 0.0005341300275176764
step: 100, loss: 0.00063440051798805594
step: 110, loss: 0.0006061287131160498
```

```
step: 120, loss: 0.00016337331908289343
step: 130, loss: 0.022158082574605942
step: 140, loss: 0.0007116972119547427
step: 150, loss: 0.006569413002580404
step: 160, loss: 0.00011440774687798694
step: 170, loss: 0.0015737803187221289
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 180, loss: 0.008264591917395592
step: 190, loss: 0.00019288503972347826
step: 200, loss: 0.0017002446111291647
step: 210, loss: 9.237071935785934e-05
step: 220, loss: 0.0008070424664765596
step: 230, loss: 0.008704615756869316
/usr/local/lib/python3.7/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=0.9853438556933484, f1=0.9831271091113611,
best_f1=0.9787709497206705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02389601431787014
step: 10, loss: 0.02943493239581585
step: 20, loss: 0.0003742764238268137
step: 30, loss: 6.935307465028018e-05
step: 40, loss: 0.001273836474865675
step: 50, loss: 0.0027690748684108257
step: 60, loss: 0.00012811111810151488
step: 70, loss: 0.09472363442182541
step: 80, loss: 0.00023022400273475796
step: 90, loss: 0.004934086930067425
step: 100, loss: 0.0077994372695568443
step: 110, loss: 0.020967252552509308
step: 120, loss: 0.01877199113368988
step: 130, loss: 8.500395779265091e-05
step: 140, loss: 0.02824929915368557
step: 150, loss: 0.00017717547598294914
step: 160, loss: 0.000907570356503129
step: 170, loss: 0.00010556101187830791
step: 180, loss: 0.010178900323808193
step: 190, loss: 0.00013173067418392748
step: 200, loss: 0.00010959129576804116
step: 210, loss: 9.840537677519023e-05
step: 220, loss: 0.038988444954156876
step: 230, loss: 0.004060945939272642
/usr/local/lib/python3.7/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=0.9830124575311437, f1=0.9784824462061155,
best_f1=0.9787709497206705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.787100912537426e-05
step: 10, loss: 0.0020388918928802013
step: 20, loss: 7.65846561989747e-05
step: 30, loss: 0.00064472743542866492
step: 40, loss: 4.8378024075645953e-05
step: 50, loss: 0.0001115609411499463
step: 60, loss: 0.03025781363248825
step: 70, loss: 0.00060143304290249994
step: 80, loss: 0.018413817510008812
step: 90, loss: 0.00035067310091108084
step: 100, loss: 0.0003301170654594898
step: 110, loss: 0.00016874627908691764
step: 120, loss: 0.002539839828386903
step: 130, loss: 0.0010128712747246027
step: 140, loss: 0.0029391315765678883
step: 150, loss: 9.604799561202526e-05
step: 160, loss: 0.00016669811157044023
step: 170, loss: 8.262037590611726e-05
step: 180, loss: 0.00029904727125540376
step: 190, loss: 6.421534635592252e-05
step: 200, loss: 3.717754952958785e-05
step: 210, loss: 5.514571603271179e-05
step: 220, loss: 0.00018171661940868944
step: 230, loss: 0.002109578112140298
/usr/local/lib/python3.7/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=0.9818594104308391, f1=0.9785310734463276,
best_f1=0.9787709497206705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.386788638541475e-05
step: 10, loss: 0.0001688938937149942
step: 20, loss: 0.0001257002295460552
step: 30, loss: 0.00013842158659826964
step: 40, loss: 0.0001672990620136261
step: 50, loss: 0.00021341574029065669
step: 60, loss: 0.0011870062444359064
step: 70, loss: 0.0011902874102815986
step: 80, loss: 0.0003540499892551452
step: 90, loss: 5.8983372582588345e-05
step: 100, loss: 0.00011884675041073933
```

```
step: 110, loss: 5.463715933728963e-05
step: 120, loss: 0.00011002954852301627
step: 130, loss: 7.87783064879477e-05
step: 140, loss: 5.2553928981069475e-05
step: 150, loss: 4.195306246401742e-05
step: 160, loss: 5.808870992041193e-05
step: 170, loss: 0.00029537672526203096
step: 180, loss: 0.013506170362234116
step: 190, loss: 0.00021055570687167346
step: 200, loss: 0.0011330224806442857
step: 210, loss: 9.566851804265752e-05
step: 220, loss: 0.00010917321196757257
step: 230, loss: 0.0003190433490090072
/usr/local/lib/python3.7/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=0.9830124575311437, f1=0.9819004524886877,
best_f1=0.9787709497206705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004148273728787899
step: 10, loss: 0.0001497698831371963
step: 20, loss: 4.162497498327866e-05
step: 30, loss: 0.00015496984997298568
step: 40, loss: 4.0243074181489646e-05
step: 50, loss: 3.357150126248598e-05
step: 60, loss: 0.0003178968618158251
step: 70, loss: 0.00010185168503085151
step: 80, loss: 9.608642722014338e-05
step: 90, loss: 5.105170566821471e-05
step: 100, loss: 3.504647975205444e-05
step: 110, loss: 0.006455290596932173
step: 120, loss: 8.956350211519748e-05
step: 130, loss: 7.179647946031764e-05
step: 140, loss: 5.432657417259179e-05
step: 150, loss: 0.01566510833799839
step: 160, loss: 0.02319229021668434
step: 170, loss: 4.6622393711004406e-05
step: 180, loss: 0.004579337779432535
step: 190, loss: 0.005496906116604805
step: 200, loss: 0.0013445193180814385
step: 210, loss: 0.0003127802337985486
step: 220, loss: 0.0004220347909722477
step: 230, loss: 4.814341809833422e-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=0.9830890642615557, f1=0.9842342342342343,
best_f1=0.9787709497206705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.364717096905224e-05
step: 10, loss: 6.500260496977717e-05
step: 20, loss: 8.992523362394422e-05
step: 30, loss: 0.00011937213275814429
step: 40, loss: 0.00018832804926205426
step: 50, loss: 5.730298289563507e-05
step: 60, loss: 0.00014439782535191625
step: 70, loss: 4.895472739008255e-05
step: 80, loss: 0.0001035359309753403
step: 90, loss: 0.002075198804959655
step: 100, loss: 0.0010406570509076118
step: 110, loss: 0.0005021105171181262
step: 120, loss: 0.00012986754882149398
step: 130, loss: 0.00039804569678381085
step: 140, loss: 0.0006864068564027548
step: 150, loss: 5.4916956287343055e-05
step: 160, loss: 5.2256153139751405e-05
step: 170, loss: 0.00031555569148622453
step: 180, loss: 6.060591113055125e-05
step: 190, loss: 8.561387221561745e-05
step: 200, loss: 3.932986146537587e-05
step: 210, loss: 0.00010319057037122548
step: 220, loss: 5.6058452173601836e-05
step: 230, loss: 4.372081093606539e-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=0.979591836734694, f1=0.979591836734694,
best_f1=0.9787709497206705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.1273269491503015e-05
step: 10, loss: 4.674652882385999e-05
step: 20, loss: 0.0006955490098334849
step: 30, loss: 0.00318462448194623
step: 40, loss: 3.110902616754174e-05
step: 50, loss: 0.0001327982172369957
step: 60, loss: 3.789171751122922e-05
step: 70, loss: 6.191518332343549e-05
step: 80, loss: 0.0001208440589834936
step: 90, loss: 0.00013620003301184624
step: 100, loss: 0.006050024647265673
```

```
step: 110, loss: 7.003240898484364e-05
step: 120, loss: 0.00024252268485724926
step: 130, loss: 5.79253064643126e-05
step: 140, loss: 5.2438732382142916e-05
step: 150, loss: 8.389751019421965e-05
step: 160, loss: 0.000233553015277721
step: 170, loss: 5.355094253900461e-05
step: 180, loss: 8.380204235436395e-05
step: 190, loss: 0.001048663747496903
step: 200, loss: 6.641069194301963e-05
step: 210, loss: 0.01276197750121355
step: 220, loss: 6.215355097083375e-05
step: 230, loss: 8.423120016232133e-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=0.9831271091113611,
best_f1=0.9787709497206705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.567165448563173e-05
step: 10, loss: 0.0001466608082409948
step: 20, loss: 0.0001033005173667334
step: 30, loss: 0.005044145509600639
step: 40, loss: 7.179694512160495e-05
step: 50, loss: 6.110193498898298e-05
step: 60, loss: 0.00010015811858465895
step: 70, loss: 8.597768464824185e-05
step: 80, loss: 0.0015853658551350236
step: 90, loss: 4.277910556993447e-05
step: 100, loss: 5.020725438953377e-05
step: 110, loss: 0.002371269278228283
step: 120, loss: 8.852902101352811e-05
step: 130, loss: 0.0003695694904308766
step: 140, loss: 0.002487187972292304
step: 150, loss: 7.281934813363478e-05
step: 160, loss: 0.0003065481723751873
step: 170, loss: 3.489748633001e-05
step: 180, loss: 0.0001078466375474818
step: 190, loss: 0.0017682835459709167
step: 200, loss: 5.518615580513142e-05
step: 210, loss: 0.0021711394656449556
step: 220, loss: 0.0001579971140017733
step: 230, loss: 0.0001689507334958762
/usr/local/lib/python3.7/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=0.983050847457627, f1=0.9830890642615557,
best_f1=0.9787709497206705
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 [00:09, 272.10it/s]
load_f1 = 0.9853438556933484
real_f1 = 0.9865168539325843
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 [00:12, 346.14it/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 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: 0.8015667200088501
step: 10, loss: 0.4245116710662842
step: 20, loss: 0.5043843984603882
step: 30, loss: 0.44183549284935
step: 40, loss: 0.40576881170272827
step: 50, loss: 0.2385338693857193
step: 60, loss: 0.15998029708862305
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 70, loss: 0.14800530672073364
step: 80, loss: 0.23028680682182312
step: 90, loss: 0.1401429921388626
step: 100, loss: 0.2442757785320282
step: 110, loss: 0.07865546643733978
step: 120, loss: 0.0318419449031353
step: 130, loss: 0.016656488180160522
step: 140, loss: 0.3262074589729309
step: 150, loss: 0.04962703958153725
step: 160, loss: 0.10209318995475769
step: 170, loss: 0.2988539934158325
step: 180, loss: 0.10017145425081253
step: 190, loss: 0.04412974789738655
step: 200, loss: 0.15649202466011047
step: 210, loss: 0.09608152508735657
step: 220, loss: 0.0802222490310669
step: 230, loss: 0.10715154558420181
step: 240, loss: 0.14955587685108185
step: 250, loss: 0.0634130984544754
step: 260, loss: 0.04000448063015938
```

```
step: 270, loss: 0.02644999697804451
step: 280, loss: 0.11534716933965683
step: 290, loss: 0.06179410591721535
step: 300, loss: 0.031001001596450806
step: 310, loss: 0.045077238231897354
step: 320, loss: 0.06167321279644966
step: 330, loss: 0.10853125900030136
step: 340, loss: 0.20690959692001343
step: 350, loss: 0.03998173400759697
step: 360, loss: 0.060377806425094604
step: 370, loss: 0.18015219271183014
step: 380, loss: 0.21724484860897064
step: 390, loss: 0.13123254477977753
step: 400, loss: 0.01576877012848854
step: 410, loss: 0.040998347103595734
step: 420, loss: 0.03226934000849724
step: 430, loss: 0.03784261271357536
step: 440, loss: 0.16141420602798462
step: 450, loss: 0.01734975539147854
step: 460, loss: 0.06150376796722412
step: 470, loss: 0.214121475815773
step: 480, loss: 0.24052317440509796
step: 490, loss: 0.0698133334517479
step: 500, loss: 0.018707210198044777
step: 510, loss: 0.1368136703968048
step: 520, loss: 0.19074510037899017
step: 530, loss: 0.14259684085845947
/usr/local/lib/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=0.9142595139844107, f1=0.9044117647058824,
best_f1=0.9044117647058824
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1359693557024002
step: 10, loss: 0.19246365129947662
step: 20, loss: 0.10602911561727524
step: 30, loss: 0.09405752271413803
step: 40, loss: 0.02197737619280815
step: 50, loss: 0.09839288145303726
step: 60, loss: 0.13133634626865387
step: 70, loss: 0.10007864981889725
step: 80, loss: 0.018588539212942123
```

```
step: 90, loss: 0.0593419224023819
step: 100, loss: 0.38451117277145386
step: 110, loss: 0.08232390135526657
step: 120, loss: 0.03946472704410553
step: 130, loss: 0.10736329853534698
step: 140, loss: 0.03435583785176277
step: 150, loss: 0.016699740663170815
step: 160, loss: 0.024875732138752937
step: 170, loss: 0.11051937937736511
step: 180, loss: 0.018092673271894455
step: 190, loss: 0.027690401300787926
step: 200, loss: 0.07356460392475128
step: 210, loss: 0.03352132812142372
step: 220, loss: 0.08784348517656326
step: 230, loss: 0.05974650755524635
step: 240, loss: 0.14935246109962463
step: 250, loss: 0.0650070458650589
step: 260, loss: 0.03796727955341339
step: 270, loss: 0.11789517849683762
step: 280, loss: 0.36083105206489563
step: 290, loss: 0.11030752956867218
step: 300, loss: 0.03734375536441803
step: 310, loss: 0.04312943294644356
step: 320, loss: 0.14546145498752594
step: 330, loss: 0.054473232477903366
step: 340, loss: 0.04749566316604614
step: 350, loss: 0.03350800648331642
step: 360, loss: 0.05316103622317314
step: 370, loss: 0.02568092755973339
step: 380, loss: 0.0591459684073925
step: 390, loss: 0.029894724488258362
step: 400, loss: 0.058084357529878616
step: 410, loss: 0.028158191591501236
step: 420, loss: 0.06818503886461258
step: 430, loss: 0.044894762337207794
step: 440, loss: 0.01891569048166275
step: 450, loss: 0.03358433023095131
step: 460, loss: 0.25303247570991516
step: 470, loss: 0.051986098289489746
step: 480, loss: 0.2552454471588135
step: 490, loss: 0.03502319008111954
step: 500, loss: 0.050896938890218735
step: 510, loss: 0.08959081023931503
step: 520, loss: 0.013109653256833553
step: 530, loss: 0.11964677274227142
/usr/local/lib/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=0.923572744014733, f1=0.9155844155844156,
best_f1=0.9155844155844156
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028351791203022003
step: 10, loss: 0.05134815722703934
step: 20, loss: 0.19107232987880707
step: 30, loss: 0.1780935376882553
step: 40, loss: 0.0026059432420879602
step: 50, loss: 0.021471478044986725
step: 60, loss: 0.009290659800171852
step: 70, loss: 0.017184700816869736
step: 80, loss: 0.014396071434020996
step: 90, loss: 0.1056656539440155
step: 100, loss: 0.0048373122699558735
step: 110, loss: 0.0468754842877388
step: 120, loss: 0.028183717280626297
step: 130, loss: 0.00990631990134716
step: 140, loss: 0.009924142621457577
step: 150, loss: 0.011164451949298382
step: 160, loss: 0.009915201924741268
step: 170, loss: 0.012840352952480316
step: 180, loss: 0.03931277617812157
step: 190, loss: 0.018874455243349075
step: 200, loss: 0.041971489787101746
step: 210, loss: 0.026273958384990692
step: 220, loss: 0.03902415558695793
step: 230, loss: 0.03400124981999397
step: 240, loss: 0.05769461393356323
step: 250, loss: 0.005623606499284506
step: 260, loss: 0.01773434318602085
step: 270, loss: 0.006370311602950096
step: 280, loss: 0.02017037384212017
step: 290, loss: 0.20736080408096313
step: 300, loss: 0.06554463505744934
step: 310, loss: 0.0730790048376617
step: 320, loss: 0.15677522122859955
step: 330, loss: 0.005988012533634901
step: 340, loss: 0.023090779781341553
step: 350, loss: 0.028118465095758438
step: 360, loss: 0.008846660144627094
step: 370, loss: 0.006954170297831297
step: 380, loss: 0.004086840897798538

```
step: 390, loss: 0.03422723710536957
step: 400, loss: 0.020240789279341698
step: 410, loss: 0.041734468191862106
step: 420, loss: 0.04509502649307251
step: 430, loss: 0.10711316019296646
step: 440, loss: 0.005744165740907192
step: 450, loss: 0.01330589409917593
step: 460, loss: 0.07688966393470764
step: 470, loss: 0.004762376192957163
step: 480, loss: 0.010461991652846336
step: 490, loss: 0.015020825900137424
step: 500, loss: 0.11051926761865616
step: 510, loss: 0.010309902019798756
step: 520, loss: 0.01588498242199421
step: 530, loss: 0.02301841601729393
/usr/local/lib/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=0.9281352747768906, f1=0.9161168708765316,
best_f1=0.9161168708765316
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009477779269218445
step: 10, loss: 0.05235464870929718
step: 20, loss: 0.2502966821193695
step: 30, loss: 0.10247215628623962
step: 40, loss: 0.022653060033917427
step: 50, loss: 0.01452564261853695
step: 60, loss: 0.022627312690019608
step: 70, loss: 0.006855448707938194
step: 80, loss: 0.14186714589595795
step: 90, loss: 0.062938705086670807
step: 100, loss: 0.012741980142891407
step: 110, loss: 0.006402525585144758
step: 120, loss: 0.0017097328091040254
step: 130, loss: 0.11945480108261108
step: 140, loss: 0.0541030615568161
step: 150, loss: 0.002350171562284231
step: 160, loss: 0.0070672426372766495
step: 170, loss: 0.0020322867203503847
step: 180, loss: 0.01818265952169895
step: 190, loss: 0.006101428065448999
step: 200, loss: 0.008639111183583736
```

```
step: 210, loss: 0.034121155738830566
step: 220, loss: 0.06757578998804092
step: 230, loss: 0.18391503393650055
step: 240, loss: 0.004121469799429178
step: 250, loss: 0.01482638530433178
step: 260, loss: 0.11546559631824493
step: 270, loss: 0.06947915256023407
step: 280, loss: 0.004710631910711527
step: 290, loss: 0.06982859969139099
step: 300, loss: 0.0014896439388394356
step: 310, loss: 0.01675558090209961
step: 320, loss: 0.01957525685429573
step: 330, loss: 0.006693301722407341
step: 340, loss: 0.015224747359752655
step: 350, loss: 0.0016081753419712186
step: 360, loss: 0.16811062395572662
step: 370, loss: 0.020101919770240784
step: 380, loss: 0.009893400594592094
step: 390, loss: 0.10826253890991211
step: 400, loss: 0.013707120902836323
step: 410, loss: 0.052062734961509705
step: 420, loss: 0.04351415857672691
step: 430, loss: 0.016369346529245377
step: 440, loss: 0.0844235047698021
step: 450, loss: 0.015392085537314415
step: 460, loss: 0.001723055844195187
step: 470, loss: 0.012672500684857368
step: 480, loss: 0.0009035262046381831
step: 490, loss: 0.04232867434620857
step: 500, loss: 0.0041958969086408615
step: 510, loss: 0.04533088952302933
step: 520, loss: 0.06829027086496353
step: 530, loss: 0.0030923879239708185
/usr/local/lib/python3.7/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=0.9249884951679707, f1=0.9165514061779623,
best_f1=0.9161168708765316
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00796132255345583
step: 10, loss: 0.009685664437711239
step: 20, loss: 0.06364756077528
step: 30, loss: 0.010674604214727879
step: 40, loss: 0.04457303136587143
step: 50, loss: 0.07066263258457184
step: 60, loss: 0.003962347283959389
```

```
step: 70, loss: 0.033174436539411545
step: 80, loss: 0.005216405261307955
step: 90, loss: 0.010612856596708298
step: 100, loss: 0.0009240704821422696
step: 110, loss: 0.013232692144811153
step: 120, loss: 0.028784800320863724
step: 130, loss: 0.0021527856588363647
step: 140, loss: 0.0004978786455467343
step: 150, loss: 0.0009888522326946259
step: 160, loss: 0.02240937016904354
step: 170, loss: 0.0564551055431366
step: 180, loss: 0.050612155348062515
step: 190, loss: 0.006206202320754528
step: 200, loss: 0.009933043271303177
step: 210, loss: 0.13119781017303467
step: 220, loss: 0.0008191318484023213
step: 230, loss: 0.0013283203588798642
step: 240, loss: 0.02601533941924572
step: 250, loss: 0.0006043638568371534
step: 260, loss: 0.05585172772407532
step: 270, loss: 0.004159099422395229
step: 280, loss: 0.0007681847200728953
step: 290, loss: 0.0002539778361096978
step: 300, loss: 0.08312467485666275
step: 310, loss: 0.017241567373275757
step: 320, loss: 0.007820679806172848
step: 330, loss: 0.002201039344072342
step: 340, loss: 0.04678363725543022
step: 350, loss: 0.04273560270667076
step: 360, loss: 0.13230028748512268
step: 370, loss: 0.01787794567644596
step: 380, loss: 0.1540345400571823
step: 390, loss: 0.0004243025614414364
step: 400, loss: 0.15845708549022675
step: 410, loss: 0.0010510949650779366
step: 420, loss: 0.0004508672282099724
step: 430, loss: 0.029045680537819862
step: 440, loss: 0.012943107634782791
step: 450, loss: 0.010155675932765007
step: 460, loss: 0.000698982912581414
step: 470, loss: 0.007965097203850746
step: 480, loss: 0.09727401286363602
step: 490, loss: 0.005695058032870293
step: 500, loss: 0.020243538543581963
step: 510, loss: 0.007085119374096394
step: 520, loss: 0.002139659132808447
step: 530, loss: 0.0014114914229139686
/usr/local/lib/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=0.9281767955801105, f1=0.9114510894761243,
best_f1=0.9114510894761243
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15038307011127472
step: 10, loss: 0.001165265217423439
step: 20, loss: 0.0008252296829596162
step: 30, loss: 0.008977300487458706
step: 40, loss: 0.0012349433964118361
step: 50, loss: 0.0016944772796705365
step: 60, loss: 0.00044741149758920074
step: 70, loss: 0.006488968152552843
step: 80, loss: 0.08574222028255463
step: 90, loss: 0.012256826274096966
step: 100, loss: 0.02471165545284748
step: 110, loss: 0.04294999688863754
step: 120, loss: 0.0043440936133265495
step: 130, loss: 0.08184251934289932
step: 140, loss: 0.001021247822791338
step: 150, loss: 0.001429409603588283
step: 160, loss: 0.0006605766830034554
step: 170, loss: 0.0016063486691564322
step: 180, loss: 0.0007998758228495717
step: 190, loss: 0.07168320566415787
step: 200, loss: 0.0004605120630003512
step: 210, loss: 0.0014727534726262093
step: 220, loss: 0.0062821959145367146
step: 230, loss: 0.0004305729235056788
step: 240, loss: 0.003352843225002289
step: 250, loss: 0.0016777863493189216
step: 260, loss: 0.0009035097318701446
step: 270, loss: 0.012233692221343517
step: 280, loss: 0.010590325109660625
step: 290, loss: 0.018180498853325844
step: 300, loss: 0.005039677955210209
step: 310, loss: 0.00024566191132180393
step: 320, loss: 0.0028637622017413378
step: 330, loss: 0.0031870370730757713
step: 340, loss: 0.022710029035806656
step: 350, loss: 0.0006727059953846037
step: 360, loss: 0.010210376232862473

```
step: 370, loss: 0.0010912916623055935
step: 380, loss: 0.02155831642448902
step: 390, loss: 0.029849577695131302
step: 400, loss: 0.00032811611890792847
step: 410, loss: 0.04311278834939003
step: 420, loss: 0.007023978047072887
step: 430, loss: 0.006332701072096825
step: 440, loss: 0.10495799779891968
step: 450, loss: 0.006541816052049398
step: 460, loss: 0.06044391915202141
step: 470, loss: 0.06312256306409836
step: 480, loss: 0.003693679813295603
step: 490, loss: 0.0009794803336262703
step: 500, loss: 0.0019127081613987684
step: 510, loss: 0.00018513885152060539
step: 520, loss: 0.020296631380915642
step: 530, loss: 0.008809404447674751
/usr/local/lib/python3.7/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=0.9281716417910447, f1=0.9170549860205033,
best_f1=0.9114510894761243
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03327455371618271
step: 10, loss: 0.05595043674111366
step: 20, loss: 0.003283220808953047
step: 30, loss: 0.030949832871556282
step: 40, loss: 0.003198351478204131
step: 50, loss: 0.0012301558163017035
step: 60, loss: 0.010834441520273685
step: 70, loss: 0.0010404759086668491
step: 80, loss: 0.016613628715276718
step: 90, loss: 0.009488148614764214
step: 100, loss: 0.00421000923961401
step: 110, loss: 0.0076150186359882355
step: 120, loss: 0.04786519333720207
step: 130, loss: 0.010929906740784645
step: 140, loss: 0.03950198367238045
step: 150, loss: 0.0006213728338479996
step: 160, loss: 0.007322150282561779
step: 170, loss: 0.0024754523765295744
step: 180, loss: 0.0012777477968484163
step: 190, loss: 0.07432691752910614
step: 200, loss: 0.005092257633805275
step: 210, loss: 0.014666866511106491
step: 220, loss: 0.0007234218064695597
```

```
step: 230, loss: 0.0039019552059471607
step: 240, loss: 0.009905637241899967
step: 250, loss: 0.002122083678841591
step: 260, loss: 0.003739273641258478
step: 270, loss: 0.00122837966773659
step: 280, loss: 0.04789876192808151
step: 290, loss: 0.02847459726035595
step: 300, loss: 0.0006462793680839241
step: 310, loss: 0.0008283652714453638
step: 320, loss: 0.12467775493860245
step: 330, loss: 0.00395313510671258
step: 340, loss: 0.002077456098049879
step: 350, loss: 0.004768833518028259
step: 360, loss: 0.00204449868761003
step: 370, loss: 0.01531816553324461
step: 380, loss: 0.00041164984577335417
step: 390, loss: 0.0007488824776373804
step: 400, loss: 0.0014232592657208443
step: 410, loss: 0.0012283164542168379
step: 420, loss: 0.0018588260281831026
step: 430, loss: 0.0003423840389586985
step: 440, loss: 0.0017919158563017845
step: 450, loss: 0.012195257470011711
step: 460, loss: 0.0014062999980524182
step: 470, loss: 0.00239367363974452
step: 480, loss: 0.0568888895213604
step: 490, loss: 0.0001782646868377924
step: 500, loss: 9.171860438073054e-05
step: 510, loss: 0.00046144158113747835
step: 520, loss: 0.00015441508730873466
step: 530, loss: 0.0005124695017002523
/usr/local/lib/python3.7/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=0.9234382339126351, f1=0.916431924882629,
best_f1=0.9114510894761243
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010303949005901814
step: 10, loss: 0.000558219850063324
step: 20, loss: 0.00019710000196937472
step: 30, loss: 0.20104198157787323
step: 40, loss: 0.0004983054241165519
step: 50, loss: 0.0009798837127164006
step: 60, loss: 0.003270606277510524
step: 70, loss: 0.0005120913847349584
step: 80, loss: 6.67582280584611e-05
```

```
step: 90, loss: 0.004771708976477385
step: 100, loss: 0.0024960446171462536
step: 110, loss: 0.005138806998729706
step: 120, loss: 0.08984946459531784
step: 130, loss: 0.0027450823690742254
step: 140, loss: 0.0005544887972064316
step: 150, loss: 0.0056360578164458275
step: 160, loss: 0.0031668555457144976
step: 170, loss: 0.02062414027750492
step: 180, loss: 0.00027074251556769013
step: 190, loss: 0.0014170801732689142
step: 200, loss: 0.004657571204006672
step: 210, loss: 0.0016768588684499264
step: 220, loss: 0.004353007301688194
step: 230, loss: 0.001701596425846219
step: 240, loss: 0.0006697785574942827
step: 250, loss: 0.012666840106248856
step: 260, loss: 0.0036889761686325073
step: 270, loss: 0.0013170785969123244
step: 280, loss: 0.0007628198945894837
step: 290, loss: 0.005316024646162987
step: 300, loss: 0.00059369090013206
step: 310, loss: 0.02494068071246147
step: 320, loss: 0.047494735568761826
step: 330, loss: 0.001087736221961677
step: 340, loss: 0.0009190603159368038
step: 350, loss: 0.009944352321326733
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 360, loss: 0.1946284919977188
step: 370, loss: 0.014681749045848846
step: 380, loss: 0.028475536033511162
step: 390, loss: 0.003190387738868594
step: 400, loss: 0.004146094899624586
step: 410, loss: 0.0014030501479282975
step: 420, loss: 0.0006409749621525407
step: 430, loss: 0.0016664230497553945
step: 440, loss: 0.002945001469925046
step: 450, loss: 0.003763692919164896
step: 460, loss: 0.0016601955285295844
step: 470, loss: 0.0004351156239863485
step: 480, loss: 0.000202003720914945
step: 490, loss: 0.00017759237380232662
step: 500, loss: 0.00048671686090528965
step: 510, loss: 0.0011829353170469403
step: 520, loss: 0.00028436697903089225
step: 530, loss: 6.315596692729741e-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 8: dev_f1=0.9251386321626617, f1=0.91728624535316,
best_f1=0.9114510894761243
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01261077355593443
step: 10, loss: 0.0037926342338323593
step: 20, loss: 0.0007498866762034595
step: 30, loss: 6.151803972898051e-05
step: 40, loss: 0.003898181254044175
step: 50, loss: 8.956932288128883e-05
step: 60, loss: 0.0021003992296755314
step: 70, loss: 0.1084495410323143
step: 80, loss: 0.0001262374862562865
step: 90, loss: 9.46138970903121e-05
step: 100, loss: 0.00038230372592806816
step: 110, loss: 0.004608051385730505
step: 120, loss: 5.962361683486961e-05
step: 130, loss: 0.0029626907780766487
step: 140, loss: 0.00068330758949974172
step: 150, loss: 0.0008996672695502639
step: 160, loss: 0.00021111407841090113
step: 170, loss: 0.000118814772577025
step: 180, loss: 9.216456237481907e-05
step: 190, loss: 0.08404608070850372
step: 200, loss: 0.07902003824710846
step: 210, loss: 0.0011556087993085384
step: 220, loss: 0.00194967957213521
step: 230, loss: 0.004934264812618494
step: 240, loss: 0.040606603026390076
step: 250, loss: 0.0008854198968037963
step: 260, loss: 0.00025005757925100625
step: 270, loss: 0.001890328130684793
step: 280, loss: 0.00019889477698598057
step: 290, loss: 0.00012587540550157428
step: 300, loss: 0.0017060566460713744
step: 310, loss: 0.0003467259812168777
step: 320, loss: 0.0011151632061228156
step: 330, loss: 0.0013525265967473388
step: 340, loss: 0.0007990918820723891
step: 350, loss: 0.0018274910980835557
step: 360, loss: 0.007513275370001793
step: 370, loss: 0.002310103503987193
step: 380, loss: 0.00047891208669170737
step: 390, loss: 0.001467556576244533
step: 400, loss: 0.0013861667830497026
step: 410, loss: 0.002627784386277199
```

```
step: 420, loss: 0.0012984235072508454
step: 430, loss: 0.0002816634369082749
step: 440, loss: 0.00018943533359561116
step: 450, loss: 0.0002365211839787662
step: 460, loss: 0.00013938694610260427
step: 470, loss: 0.09635601192712784
step: 480, loss: 0.11625936627388
step: 490, loss: 0.003486744826659956
step: 500, loss: 0.0051974705420434475
step: 510, loss: 0.004325024783611298
step: 520, loss: 0.0005542718572542071
step: 530, loss: 0.0004813319246750325
/usr/local/lib/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=0.9288040949278734, f1=0.9166278528178855,
best_f1=0.9166278528178855
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0025294513907283545
step: 10, loss: 0.00029688897169921547
step: 20, loss: 0.00011980170529568568
step: 30, loss: 0.0011653706897050142
step: 40, loss: 0.003954081330448389
step: 50, loss: 0.00028495557489804924
step: 60, loss: 0.003955611493438482
step: 70, loss: 0.0014479950768873096
step: 80, loss: 0.001658673514612019
step: 90, loss: 0.00012074101687176153
step: 100, loss: 0.007174021098762751
step: 110, loss: 0.0038615260273218155
step: 120, loss: 0.002640962600708008
step: 130, loss: 0.0005656818975694478
step: 140, loss: 0.0037145509850233793
step: 150, loss: 0.0048587266355752945
step: 160, loss: 0.0005769329727627337
step: 170, loss: 9.745104762259871e-05
step: 180, loss: 0.010171647183597088
step: 190, loss: 0.0013932815054431558
step: 200, loss: 0.0013744478346779943
step: 210, loss: 0.00015038950368762016
step: 220, loss: 0.0030497321859002113
step: 230, loss: 0.001321746502071619
```

```
step: 240, loss: 0.00019284260633867234
step: 250, loss: 0.00039641649345867336
step: 260, loss: 0.0016071362188085914
step: 270, loss: 0.010174287483096123
step: 280, loss: 0.00027302379021421075
step: 290, loss: 4.118186188861728e-05
step: 300, loss: 0.00011797467595897615
step: 310, loss: 0.10749126225709915
step: 320, loss: 0.00010532219312153757
step: 330, loss: 0.10833537578582764
step: 340, loss: 0.015486649237573147
step: 350, loss: 0.00013898943143431097
step: 360, loss: 0.004699924029409885
step: 370, loss: 0.001653241808526218
step: 380, loss: 0.0005041435360908508
step: 390, loss: 0.0008896688814274967
step: 400, loss: 0.0024159278254956007
step: 410, loss: 0.020354019477963448
step: 420, loss: 0.0011916577350348234
step: 430, loss: 7.531550363637507e-05
step: 440, loss: 2.6169644115725532e-05
step: 450, loss: 0.007318742573261261
step: 460, loss: 0.00019353015522938222
step: 470, loss: 4.318505307310261e-05
step: 480, loss: 0.0004301457083784044
step: 490, loss: 0.060968395322561264
step: 500, loss: 0.0022412666585296392
step: 510, loss: 0.0020813976880162954
step: 520, loss: 0.0007077055051922798
step: 530, loss: 0.007500726263970137
/usr/local/lib/python3.7/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=0.9255121042830541, f1=0.9161592505854801,
best_f1=0.9166278528178855
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006220668437890708
step: 10, loss: 0.00041801013867370784
step: 20, loss: 0.00029465986881405115
step: 30, loss: 0.006916152313351631
step: 40, loss: 8.408150461036712e-05
step: 50, loss: 0.0004685607273131609
step: 60, loss: 4.427144813234918e-05
step: 70, loss: 7.115741755114868e-05
step: 80, loss: 3.8217243854887784e-05
step: 90, loss: 0.00029106796137057245
```

```
step: 100, loss: 2.9261171221151017e-05
step: 110, loss: 3.068412479478866e-05
step: 120, loss: 0.0048766955733299255
step: 130, loss: 0.00021488437778316438
step: 140, loss: 3.644271055236459e-05
step: 150, loss: 6.87902356730774e-05
step: 160, loss: 0.0009055727859959006
step: 170, loss: 0.0001166757574537769
step: 180, loss: 3.5634129744721577e-05
step: 190, loss: 0.00579194026067853
step: 200, loss: 0.000352477771230042
step: 210, loss: 5.2370134653756395e-05
step: 220, loss: 3.4237775253131986e-05
step: 230, loss: 0.0015675489557906985
step: 240, loss: 3.461341839283705e-05
step: 250, loss: 0.0012960307067260146
step: 260, loss: 2.5715238734846935e-05
step: 270, loss: 0.01154785230755806
step: 280, loss: 0.003325813915580511
step: 290, loss: 0.0033915089443325996
step: 300, loss: 0.018734030425548553
step: 310, loss: 0.022124765440821648
step: 320, loss: 0.007946545258164406
step: 330, loss: 0.0038775356952100992
step: 340, loss: 0.0001602169795660302
step: 350, loss: 0.0004106271080672741
step: 360, loss: 0.00013175800268072635
step: 370, loss: 0.0013636834919452667
step: 380, loss: 7.332390669034794e-05
step: 390, loss: 0.019541259855031967
step: 400, loss: 0.0009076670394279063
step: 410, loss: 0.00026553933327586353
step: 420, loss: 0.004123051185160875
step: 430, loss: 0.0005852609756402671
step: 440, loss: 0.0012845896417275071
step: 450, loss: 0.0003391027566976845
step: 460, loss: 0.005606865510344505
step: 470, loss: 0.0036535686813294888
step: 480, loss: 0.014232795685529709
step: 490, loss: 0.005537676624953747
step: 500, loss: 0.001512342831119895
step: 510, loss: 0.00020594181842170656
step: 520, loss: 4.070020804647356e-05
step: 530, loss: 3.7862260796828195e-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 11: dev_f1=0.9303621169916435, f1=0.9192200557103064,
best_f1=0.9192200557103064
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010144623956875876
step: 10, loss: 6.560909969266504e-05
step: 20, loss: 2.2802207240602e-05
step: 30, loss: 0.00018177868332713842
step: 40, loss: 0.0005758345359936357
step: 50, loss: 0.0001989330048672855
step: 60, loss: 8.519885886926204e-05
step: 70, loss: 0.18559978902339935
step: 80, loss: 0.00034422401222400367
step: 90, loss: 0.0012242899974808097
step: 100, loss: 0.0009411762584932148
step: 110, loss: 0.00011205968621652573
step: 120, loss: 9.756865620147437e-05
step: 130, loss: 0.00012194610462756827
step: 140, loss: 5.8653909945860505e-05
step: 150, loss: 0.0001388191885780543
step: 160, loss: 0.00017023221880663186
step: 170, loss: 0.003644591197371483
step: 180, loss: 0.0006081857136450708
step: 190, loss: 2.907157249865122e-05
step: 200, loss: 0.00017877329082693905
step: 210, loss: 0.0003326285514049232
step: 220, loss: 0.00018090337107423693
step: 230, loss: 0.0024873788934201
step: 240, loss: 0.00011145596363348886
step: 250, loss: 0.02700163424015045
step: 260, loss: 0.0004674117371905595
step: 270, loss: 0.002986076520755887
step: 280, loss: 0.0004588724987115711
step: 290, loss: 0.0033680670894682407
step: 300, loss: 0.0009433917002752423
step: 310, loss: 0.00023710451205261052
step: 320, loss: 0.0018041158327832818
step: 330, loss: 0.0016673905774950981
step: 340, loss: 0.0004917646292597055
step: 350, loss: 0.011653576046228409
step: 360, loss: 0.001058473251760006
step: 370, loss: 4.258755507180467e-05
step: 380, loss: 0.00017677177675068378
step: 390, loss: 5.911331390962005e-05

```
step: 400, loss: 0.0008905114373192191
step: 410, loss: 0.0014200258301571012
step: 420, loss: 0.007922150194644928
step: 430, loss: 0.0015900626312941313
step: 440, loss: 0.003050523577257991
step: 450, loss: 0.004158497788012028
step: 460, loss: 0.0004783785261679441
step: 470, loss: 0.0025820957962423563
step: 480, loss: 0.00011020776582881808
step: 490, loss: 0.0012987037189304829
step: 500, loss: 0.006030887831002474
step: 510, loss: 0.0006589583354070783
step: 520, loss: 0.008718809112906456
step: 530, loss: 0.00012428023910615593
/usr/local/lib/python3.7/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=0.930341280972417, f1=0.9176029962546816,
best_f1=0.9192200557103064
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00038775659049861133
step: 10, loss: 4.28972089139279e-05
step: 20, loss: 0.010238029062747955
step: 30, loss: 0.025044895708560944
step: 40, loss: 0.0003781623672693968
step: 50, loss: 0.0017278059385716915
step: 60, loss: 2.8985370590817183e-05
step: 70, loss: 7.509395072702318e-05
step: 80, loss: 0.01570931263267994
step: 90, loss: 0.00028260290855541825
step: 100, loss: 4.254985469742678e-05
step: 110, loss: 0.0005403929972089827
step: 120, loss: 0.0004930479917675257
step: 130, loss: 0.0011156355030834675
step: 140, loss: 0.00011466386786196381
step: 150, loss: 5.4882068070583045e-05
step: 160, loss: 0.0005030675092712045
step: 170, loss: 0.0023742043413221836
step: 180, loss: 0.005858609918504953
step: 190, loss: 0.00016899099864531308
step: 200, loss: 0.0027928149793297052
step: 210, loss: 0.0051654730923473835
step: 220, loss: 5.089673140901141e-05
step: 230, loss: 0.0027689633425325155
step: 240, loss: 0.00034105824306607246
step: 250, loss: 0.012168553657829762
```

```
step: 260, loss: 0.030513519421219826
step: 270, loss: 0.04508942738175392
step: 280, loss: 3.2575881050433964e-05
step: 290, loss: 4.771065869135782e-05
step: 300, loss: 0.00047667371109128
step: 310, loss: 0.0005908438470214605
step: 320, loss: 5.9564765251707286e-05
step: 330, loss: 0.0001763576437951997
step: 340, loss: 3.443810055614449e-05
step: 350, loss: 0.00948311947286129
step: 360, loss: 0.0003178034967277199
step: 370, loss: 0.001699417713098228
step: 380, loss: 7.706564792897552e-05
step: 390, loss: 0.0012972032418474555
step: 400, loss: 5.2904932090314105e-05
step: 410, loss: 9.736568608786911e-05
step: 420, loss: 0.0003941325703635812
step: 430, loss: 0.0035300971940159798
step: 440, loss: 0.00026573569630272686
step: 450, loss: 0.0007141334936022758
step: 460, loss: 0.0005783337983302772
step: 470, loss: 0.019459273666143417
step: 480, loss: 0.0004226339515298605
step: 490, loss: 0.0006607035757042468
step: 500, loss: 0.0009002832230180502
step: 510, loss: 9.578536992194131e-05
step: 520, loss: 0.0008052241173572838
step: 530, loss: 0.00028572132578119636
/usr/local/lib/python3.7/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=0.9302540415704387, f1=0.9198895027624309,
best_f1=0.9192200557103064
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.035260932752863e-05
step: 10, loss: 0.00038285990012809634
step: 20, loss: 4.7015051677590236e-05
step: 30, loss: 3.992455094703473e-05
step: 40, loss: 4.8051635531010106e-05
step: 50, loss: 0.0016378412256017327
step: 60, loss: 3.2624553568894044e-05
step: 70, loss: 0.00015893766249064356
step: 80, loss: 0.0003165339876431972
step: 90, loss: 0.00011770275887101889
step: 100, loss: 0.000456042995210737
step: 110, loss: 7.878917676862329e-05
```

```
step: 120, loss: 0.00022879971947986633
step: 130, loss: 0.0002260697219753638
step: 140, loss: 0.019788380712270737
step: 150, loss: 0.00023185941972769797
step: 160, loss: 0.00025773068773560226
step: 170, loss: 0.00012045874609611928
step: 180, loss: 0.0034790271893143654
step: 190, loss: 0.00029794868896715343
step: 200, loss: 0.00033823554986156523
step: 210, loss: 0.0012372104683890939
step: 220, loss: 0.00014561951684299856
step: 230, loss: 3.05899593513459e-05
step: 240, loss: 9.915282862493768e-05
step: 250, loss: 5.504305590875447e-05
step: 260, loss: 1.846231498348061e-05
step: 270, loss: 0.02894441783428192
step: 280, loss: 4.098549106856808e-05
step: 290, loss: 0.0005679457681253552
step: 300, loss: 0.0007123675313778222
step: 310, loss: 3.6158147850073874e-05
step: 320, loss: 0.00018777011428028345
step: 330, loss: 0.00012850240455009043
step: 340, loss: 0.006388332694768906
step: 350, loss: 0.0016552676679566503
step: 360, loss: 0.0008441862883046269
step: 370, loss: 0.00025002832990139723
step: 380, loss: 0.00041012419387698174
step: 390, loss: 0.0009859324200078845
step: 400, loss: 0.0002631042734719813
step: 410, loss: 0.00015346930013038218
step: 420, loss: 9.339419921161607e-05
step: 430, loss: 0.007090008817613125
step: 440, loss: 8.398221689276397e-05
step: 450, loss: 0.0009478341671638191
step: 460, loss: 0.0003382039722055197
step: 470, loss: 0.0006954415002837777
step: 480, loss: 2.4612303604953922e-05
step: 490, loss: 4.373032061266713e-05
step: 500, loss: 7.004595681792125e-05
step: 510, loss: 3.645316246547736e-05
step: 520, loss: 5.4163851018529385e-05
step: 530, loss: 0.00012939010048285127
/usr/local/lib/python3.7/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=0.9301675977653632, f1=0.9202797202797203,
best_f1=0.9192200557103064
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017960580589715391
step: 10, loss: 4.569400698528625e-05
step: 20, loss: 9.001341823022813e-05
step: 30, loss: 5.006411447538994e-05
step: 40, loss: 2.4146867872332223e-05
step: 50, loss: 0.00011568440095288679
step: 60, loss: 3.139496402582154e-05
step: 70, loss: 3.2545838621445e-05
step: 80, loss: 6.751050386810675e-05
step: 90, loss: 1.714359677862376e-05
step: 100, loss: 1.6223479178734124e-05
step: 110, loss: 0.0017329563852399588
step: 120, loss: 0.000543405709322542
step: 130, loss: 4.670603811973706e-05
step: 140, loss: 1.8406459275865927e-05
step: 150, loss: 0.0002352682058699429
step: 160, loss: 5.4279909818433225e-05
step: 170, loss: 0.001831036526709795
step: 180, loss: 4.000199260190129e-05
step: 190, loss: 0.00026462198002263904
step: 200, loss: 8.016890205908567e-05
step: 210, loss: 0.0004230555205140263
step: 220, loss: 0.00014006311539560556
step: 230, loss: 5.2032817620784044e-05
step: 240, loss: 0.00012779617100022733
step: 250, loss: 0.00019292577053420246
step: 260, loss: 0.06448429077863693
step: 270, loss: 0.000520861882250756
step: 280, loss: 4.283722591935657e-05
step: 290, loss: 0.0012489601504057646
step: 300, loss: 0.01329758856445551
step: 310, loss: 0.00018247676780447364
step: 320, loss: 0.00011916666699107736
step: 330, loss: 0.0005737406318075955
step: 340, loss: 0.01650279574096203
step: 350, loss: 0.0004154241469223052
step: 360, loss: 5.333541048457846e-05
step: 370, loss: 6.055486301193014e-05
step: 380, loss: 0.00011557927064131945
step: 390, loss: 0.00035235314862802625
step: 400, loss: 2.7899468477698974e-05
step: 410, loss: 5.727373718400486e-05
step: 420, loss: 7.043997902655974e-05
step: 430, loss: 2.2656700821244158e-05
step: 440, loss: 0.0004109699802938849
step: 450, loss: 0.0029769763350486755
```

```
step: 460, loss: 2.2321562937577255e-05
step: 470, loss: 5.343799421098083e-05
step: 480, loss: 0.0005139539716765285
step: 490, loss: 0.0010545760160312057
step: 500, loss: 3.3090764191001654e-05
step: 510, loss: 0.017804667353630066
step: 520, loss: 7.586739957332611e-05
step: 530, loss: 2.497700188541785e-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=0.9315323707498836, f1=0.9210649229332087,
best_f1=0.9210649229332087
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.")
5742it [00:16, 345.72it/s]
```

```
load_f1 = 0.928212162780064
real_f1 = 0.9206927985414768
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 [00:12, 362.58it/s]
```

4.1.4 Amazon-Google - Running the matcher

```python
[ ]: # 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 O2: FP16 training with FP32 batchnorm and FP32 master weights.

395

```
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: 0.8300186991691589
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.06490509957075119
step: 20, loss: 0.3793472945690155
step: 30, loss: 0.3707546293735504
step: 40, loss: 0.5154072046279907
step: 50, loss: 0.30099767446517944
step: 60, loss: 0.363876610994339
step: 70, loss: 0.2597098648548126
step: 80, loss: 0.2942117154598236
step: 90, loss: 0.4492667317390442
step: 100, loss: 0.20705467462539673
step: 110, loss: 0.2908226549625397
step: 120, loss: 0.24972103536128998
step: 130, loss: 0.2180168479681015
```

```
step: 140, loss: 0.2588222324848175
step: 150, loss: 0.29451531171798706
step: 160, loss: 0.2364419400691986
step: 170, loss: 0.1679026186466217
step: 180, loss: 0.2395094335079193
step: 190, loss: 0.2017327845096588
step: 200, loss: 0.18106108903884888
step: 210, loss: 0.4780985116958618
/usr/local/lib/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=0.5047438330170777, f1=0.5140562248995983,
best_f1=0.5140562248995983
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07968694716691971
step: 10, loss: 0.09203134477138519
step: 20, loss: 0.2085149586200714
step: 30, loss: 0.15958532691001892
step: 40, loss: 0.04037568345665932
step: 50, loss: 0.23010489344596863
step: 60, loss: 0.07708368450403214
step: 70, loss: 0.22980482876300812
step: 80, loss: 0.3087536692619324
step: 90, loss: 0.17271938920021057
step: 100, loss: 0.0964159145951271
step: 110, loss: 0.08377044647932053
step: 120, loss: 0.18192780017852783
step: 130, loss: 0.290659099817276
step: 140, loss: 0.27655166387557983
step: 150, loss: 0.24177439510822296
step: 160, loss: 0.19399237632751465
step: 170, loss: 0.1761600822210312
step: 180, loss: 0.2599106431007385
step: 190, loss: 0.23673857748508453
step: 200, loss: 0.17967334389686584
step: 210, loss: 0.321493923664093
/usr/local/lib/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=0.5393939393939393, f1=0.5709923664122137,
best_f1=0.5709923664122137
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26993846893310547
step: 10, loss: 0.18995530903339386
step: 20, loss: 0.30274683237075806
step: 30, loss: 0.09917538613080978
step: 40, loss: 0.13995271921157837
step: 50, loss: 0.1298438161611557
step: 60, loss: 0.2678304612636566
step: 70, loss: 0.07632572948932648
step: 80, loss: 0.08671435713768005
step: 90, loss: 0.13791710138320923
step: 100, loss: 0.032850898802280426
step: 110, loss: 0.14143729209899902
step: 120, loss: 0.11059051752090454
step: 130, loss: 0.15138180553913116
step: 140, loss: 0.3934134244918823
step: 150, loss: 0.32159534096717834
step: 160, loss: 0.14702732861042023
step: 170, loss: 0.2884335219860077
step: 180, loss: 0.08620435744524002
step: 190, loss: 0.34467145800590515
step: 200, loss: 0.1130581870675087
step: 210, loss: 0.31858348846435547
/usr/local/lib/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=0.5656565656565656, f1=0.598326359832636,
best_f1=0.598326359832636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1669633835554123
step: 10, loss: 0.03546725586056709
step: 20, loss: 0.0774264931678772
step: 30, loss: 0.07123205065727234
step: 40, loss: 0.1607426404953003
step: 50, loss: 0.08916439116001129
step: 60, loss: 0.049063198268413544
step: 70, loss: 0.34536033868789673
```

```
step: 80, loss: 0.11989696323871613
step: 90, loss: 0.03301994875073433
step: 100, loss: 0.13063406944274902
step: 110, loss: 0.20418238639831543
step: 120, loss: 0.07369434088468552
step: 130, loss: 0.1612521857023239
step: 140, loss: 0.09961211681365967
step: 150, loss: 0.152373269200325
step: 160, loss: 0.0479377917945385
step: 170, loss: 0.06413891166448593
step: 180, loss: 0.220417320728302
step: 190, loss: 0.23010705411434174
step: 200, loss: 0.16729114949703217
step: 210, loss: 0.06537719070911407
/usr/local/lib/python3.7/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=0.5410447761194029, f1=0.5836575875486382,
best_f1=0.598326359832636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05552072077989578
step: 10, loss: 0.05159789323806763
step: 20, loss: 0.026285238564014435
step: 30, loss: 0.04301043599843979
step: 40, loss: 0.17864935100078583
step: 50, loss: 0.12020064145326614
step: 60, loss: 0.04940402880311012
step: 70, loss: 0.09353432059288025
step: 80, loss: 0.03533273562788963
step: 90, loss: 0.15523073077201843
step: 100, loss: 0.11969796568155289
step: 110, loss: 0.025293558835983276
step: 120, loss: 0.07002739608287811
step: 130, loss: 0.07067501544952393
step: 140, loss: 0.04280441999435425
step: 150, loss: 0.10650131851434708
step: 160, loss: 0.10843590646982193
step: 170, loss: 0.0929868295788765
step: 180, loss: 0.3317057192325592
step: 190, loss: 0.1636005938053131
step: 200, loss: 0.09650442749261856
step: 210, loss: 0.08571341633796692
/usr/local/lib/python3.7/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=0.5562632696390659, f1=0.5294117647058824,
```

```
best_f1=0.598326359832636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03398895263671875
step: 10, loss: 0.014302410185337067
step: 20, loss: 0.04715421423316002
step: 30, loss: 0.07271140813827515
step: 40, loss: 0.010628691874444485
step: 50, loss: 0.06344104558229446
step: 60, loss: 0.0896785706281662
step: 70, loss: 0.02666774019598961
step: 80, loss: 0.12041284888982773
step: 90, loss: 0.04539605975151062
step: 100, loss: 0.023657288402318954
step: 110, loss: 0.0044631920754909515
step: 120, loss: 0.05388474091887474
step: 130, loss: 0.030859846621751785
step: 140, loss: 0.06604894995689392
step: 150, loss: 0.09981485456228256
step: 160, loss: 0.1479898989200592
step: 170, loss: 0.028392944484949112
step: 180, loss: 0.01049861405044794
step: 190, loss: 0.09880705922842026
step: 200, loss: 0.022014427930116653
step: 210, loss: 0.017854660749435425
/usr/local/lib/python3.7/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=0.5342237061769616, f1=0.5366726296958856,
best_f1=0.598326359832636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.037441760301589966
step: 10, loss: 0.04557556286454201
step: 20, loss: 0.1910131871700287
step: 30, loss: 0.0070851026102900505
step: 40, loss: 0.0624687559902668
step: 50, loss: 0.07857085764408112
step: 60, loss: 0.20129047334194183
step: 70, loss: 0.05096055567264557
step: 80, loss: 0.022694764658808708
step: 90, loss: 0.06620750576257706
step: 100, loss: 0.010205315425992012
step: 110, loss: 0.17432227730751038
step: 120, loss: 0.16920238733291626
step: 130, loss: 0.00692064268514514
```

```
step: 140, loss: 0.00462180795148015
step: 150, loss: 0.06628444045782089
step: 160, loss: 0.09485410898923874
step: 170, loss: 0.1172945573925972
step: 180, loss: 0.0053933910094201565
step: 190, loss: 0.034586742520332336
step: 200, loss: 0.06181282922625542
step: 210, loss: 0.10885083675384521
/usr/local/lib/python3.7/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=0.5387323943661972, f1=0.5434380776340111,
best_f1=0.598326359832636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04499286785721779
step: 10, loss: 0.06697148084640503
step: 20, loss: 0.02878543734550476
step: 30, loss: 0.08322105556726456
step: 40, loss: 0.048088423907756805
step: 50, loss: 0.011278694495558739
step: 60, loss: 0.34687721729278564
step: 70, loss: 0.005873464979231358
step: 80, loss: 0.05782892182469368
step: 90, loss: 0.08598518371582031
step: 100, loss: 0.03278382495045662
step: 110, loss: 0.004817671608179808
step: 120, loss: 0.016969947144389153
step: 130, loss: 0.11461798846721649
step: 140, loss: 0.0044730291701853275
step: 150, loss: 0.005379854701459408
step: 160, loss: 0.056250639259815216
step: 170, loss: 0.10172603279352188
step: 180, loss: 0.029087521135807037
step: 190, loss: 0.042459577322006226
step: 200, loss: 0.12321585416793823
step: 210, loss: 0.06626702100038528
/usr/local/lib/python3.7/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=0.5415162454873648, f1=0.53515625, best_f1=0.598326359832636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011481206864118576
step: 10, loss: 0.020222987979650497
step: 20, loss: 0.03396158292889595
```

```
step: 30, loss: 0.17717967927455902
step: 40, loss: 0.06422212719917297
step: 50, loss: 0.008522926829755306
step: 60, loss: 0.020927706733345985
step: 70, loss: 0.01122854370623827
step: 80, loss: 0.014503237791359425
step: 90, loss: 0.08103648573160172
step: 100, loss: 0.021962661296129227
step: 110, loss: 0.010402942076325417
step: 120, loss: 0.01334876473993063
step: 130, loss: 0.013932544738054276
step: 140, loss: 0.04724960774183273
step: 150, loss: 0.046314042061567307
step: 160, loss: 0.004020704422146082
step: 170, loss: 0.053852617740631104
step: 180, loss: 0.048187535256147385
step: 190, loss: 0.08012165129184723
step: 200, loss: 0.024515429511666298
step: 210, loss: 0.10810276120901108
/usr/local/lib/python3.7/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=0.5325670498084291, f1=0.5167652859960552,
best_f1=0.598326359832636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009280656464397907
step: 10, loss: 0.012728352099657059
step: 20, loss: 0.04600915685296059
step: 30, loss: 0.008968198671936989
step: 40, loss: 0.012718599289655685
step: 50, loss: 0.004805580712854862
step: 60, loss: 0.004966500215232372
step: 70, loss: 0.06645932793617249
step: 80, loss: 0.0009302472462877631
step: 90, loss: 0.07280661910772324
step: 100, loss: 0.0010433276183903217
step: 110, loss: 0.004733223468065262
step: 120, loss: 0.004569693934172392
step: 130, loss: 0.015081661753356457
step: 140, loss: 0.0034223096445202827
step: 150, loss: 0.12628351151943207
step: 160, loss: 0.010107903741300106
step: 170, loss: 0.02257111482322216
step: 180, loss: 0.0032130880281329155
step: 190, loss: 0.14005543291568756
step: 200, loss: 0.012163562700152397
```

step: 210, loss: 0.07268618047237396
/usr/local/lib/python3.7/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=0.5361552028218695, f1=0.512241054613936,
best_f1=0.598326359832636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03531704470515251
step: 10, loss: 0.19274647533893585
step: 20, loss: 0.008807888254523277
step: 30, loss: 0.03812893107533455
step: 40, loss: 0.03041134774684906
step: 50, loss: 0.002843103113872993
step: 60, loss: 0.006837894674390554
step: 70, loss: 0.0033341296948492527
step: 80, loss: 0.07695766538381577
step: 90, loss: 0.05838211625814438
step: 100, loss: 0.031944334506988525
step: 110, loss: 0.11655104160308838
step: 120, loss: 0.006941227708011866
step: 130, loss: 0.010275272652506828
step: 140, loss: 0.02358575165271759
step: 150, loss: 0.012709718197584152
step: 160, loss: 0.0025111890863627195
step: 170, loss: 0.1674751192331314
step: 180, loss: 0.003070353763177991
step: 190, loss: 0.005247362889349461
step: 200, loss: 0.00847592018544674
step: 210, loss: 0.0029915517661720514
/usr/local/lib/python3.7/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=0.5588235294117646, f1=0.5244618395303327,
best_f1=0.598326359832636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016351185739040375
step: 10, loss: 0.008766064420342445
step: 20, loss: 0.11188706755638123
step: 30, loss: 0.004558993503451347
step: 40, loss: 0.0570489764213562
step: 50, loss: 0.011602260172367096
step: 60, loss: 0.007001896388828754
step: 70, loss: 0.00330945523455739
step: 80, loss: 0.04499336704611778

```
step: 90, loss: 0.0033001387491822243
step: 100, loss: 0.005496481899172068
step: 110, loss: 0.0019510674756020308
step: 120, loss: 0.0036378279328346252
step: 130, loss: 0.038590576499700546
step: 140, loss: 0.013951092027127743
step: 150, loss: 0.0016480067279189825
step: 160, loss: 0.003890999825671315
step: 170, loss: 0.010278104804456234
step: 180, loss: 0.00450814189389348
step: 190, loss: 0.15403547883033752
step: 200, loss: 0.004728730767965317
step: 210, loss: 0.015473463572561741
/usr/local/lib/python3.7/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=0.55, f1=0.5175983436853003, best_f1=0.598326359832636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009826000314205885
step: 10, loss: 0.011689441278576851
step: 20, loss: 0.0029035431798547506
step: 30, loss: 0.003790880087763071
step: 40, loss: 0.01586763560771942
step: 50, loss: 0.0014983332948759198
step: 60, loss: 0.008808048442006111
step: 70, loss: 0.02924743853509426
step: 80, loss: 0.012382817454636097
step: 90, loss: 0.001379061839543283
step: 100, loss: 0.060273174196481705
step: 110, loss: 0.00685111666098237
step: 120, loss: 0.0012053013779222965
step: 130, loss: 0.0022358500864356756
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 140, loss: 0.043407946825027466
step: 150, loss: 0.0017938301898539066
step: 160, loss: 0.003799030790105462
step: 170, loss: 0.16362543404102325
step: 180, loss: 0.0197137557559657097
step: 190, loss: 0.004638676065951586
step: 200, loss: 0.007092122919857502
step: 210, loss: 0.03624146804213524
/usr/local/lib/python3.7/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=0.5468164794007491, f1=0.5392354124748492,
best_f1=0.598326359832636
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002613202203065157
step: 10, loss: 0.001224431092850864
step: 20, loss: 0.0032161923591047525
step: 30, loss: 0.014092403464019299
step: 40, loss: 0.004088675137609243
step: 50, loss: 0.00277451379224658
step: 60, loss: 0.10497882217168808
step: 70, loss: 0.018673967570066452
step: 80, loss: 0.08457383513450623
step: 90, loss: 0.10485818237066269
step: 100, loss: 0.007890413515269756
step: 110, loss: 0.0053333016112446785
step: 120, loss: 0.012496245093643665
step: 130, loss: 0.005182444117963314
step: 140, loss: 0.0008582309819757938
step: 150, loss: 0.18630388379096985
step: 160, loss: 0.0014954553917050362
step: 170, loss: 0.03596971556544304
step: 180, loss: 0.06617721915245056
step: 190, loss: 0.014406138099730015
step: 200, loss: 0.004738971125334501
step: 210, loss: 0.028012415394186974
/usr/local/lib/python3.7/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=0.5492424242424243, f1=0.537987679671458,
best_f1=0.598326359832636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0064093791879713535
step: 10, loss: 0.011963452212512493
step: 20, loss: 0.0045647937804460526
step: 30, loss: 0.13447001576423645
step: 40, loss: 0.0019439348252490163
step: 50, loss: 0.008227232843637466
step: 60, loss: 0.009012722410261631
step: 70, loss: 0.0019105894025415182
step: 80, loss: 0.011920446529984474
step: 90, loss: 0.021333010867238045
step: 100, loss: 0.004339347593486309
step: 110, loss: 0.0019554675091058016
step: 120, loss: 0.2567243278026581
step: 130, loss: 0.004485683050006628
step: 140, loss: 0.0011778193293139338
```

```
step: 150, loss: 0.018359987065196037
step: 160, loss: 0.005675586871802807
step: 170, loss: 0.009169564582407475
step: 180, loss: 0.0017746298108249903
step: 190, loss: 0.05361318960785866
step: 200, loss: 0.004032694268971682
step: 210, loss: 0.029181210324168205
/usr/local/lib/python3.7/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=0.5468451242829828, f1=0.5338809034907598,
best_f1=0.598326359832636
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.")
2293it [00:04, 559.38it/s]
load_f1 = 0.5504273504273505
real_f1 = 0.544839255499154
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 [00:12, 352.79it/s]
```

4.1.5   Walmart-Amazon - Running the matcher

```python
[ ]: # 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/
```

```
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
```

407

```
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: 0.850167453289032
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.1655227690935135
step: 20, loss: 0.1512197107076645
step: 30, loss: 0.49564844369888306
step: 40, loss: 0.2476639300584793
step: 50, loss: 0.30912667512893677
step: 60, loss: 0.36445003747940063
step: 70, loss: 0.17423228919506073
step: 80, loss: 0.5051755905151367
step: 90, loss: 0.2412043958902359
step: 100, loss: 0.22281664609909058
step: 110, loss: 0.23485003411769867
step: 120, loss: 0.4133383333683014
step: 130, loss: 0.34211522340774536
step: 140, loss: 0.31376951932907104
step: 150, loss: 0.2634420692920685
step: 160, loss: 0.183495432138443
step: 170, loss: 0.3897187411785126
step: 180, loss: 0.2765911221504211

```

```
step: 190, loss: 0.13766340911388397
/usr/local/lib/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=0.5760000000000001, f1=0.6312997347480105,
best_f1=0.6312997347480105
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23637287318706512
step: 10, loss: 0.06170308589935303
step: 20, loss: 0.11776485294103622
step: 30, loss: 0.1408710926771164
step: 40, loss: 0.44085493683815
step: 50, loss: 0.28816142678260803
step: 60, loss: 0.08128667622804642
step: 70, loss: 0.23695357143878937
step: 80, loss: 0.17576546967029572
step: 90, loss: 0.059276532381772995
step: 100, loss: 0.1902966946363449
step: 110, loss: 0.15227043628692627
step: 120, loss: 0.17917503416538239
step: 130, loss: 0.09123601764440536
step: 140, loss: 0.2307703197002411
step: 150, loss: 0.03418679162859917
step: 160, loss: 0.11066599190235138
step: 170, loss: 0.31777530908584595
step: 180, loss: 0.11414851993322372
step: 190, loss: 0.1614735722541809
/usr/local/lib/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=0.7513812154696132, f1=0.7313019390581716,
best_f1=0.7313019390581716
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10755188763141632
step: 10, loss: 0.2967492640018463
step: 20, loss: 0.056679096072912216
```

```
step: 30, loss: 0.04153234139084816
step: 40, loss: 0.03963930159807205
step: 50, loss: 0.1802026480436325
step: 60, loss: 0.04795122891664505
step: 70, loss: 0.12380363047122955
step: 80, loss: 0.22019889950752258
step: 90, loss: 0.06780440360307693
step: 100, loss: 0.07254137098789215
step: 110, loss: 0.257583886384964
step: 120, loss: 0.0667315348982811
step: 130, loss: 0.017768556252121925
step: 140, loss: 0.08538202196359634
step: 150, loss: 0.17735300958156586
step: 160, loss: 0.1905943602323532
step: 170, loss: 0.02976483292877674
step: 180, loss: 0.07851861417293549
step: 190, loss: 0.1788814514875412
/usr/local/lib/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=0.7758620689655173, f1=0.7755102040816326,
best_f1=0.7755102040816326
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.199260026216500696
step: 10, loss: 0.05388612300157547
step: 20, loss: 0.015228705480694477
step: 30, loss: 0.0737965852022171
step: 40, loss: 0.0027220030315220356
step: 50, loss: 0.06980481743812561
step: 60, loss: 0.09397035837173462
step: 70, loss: 0.010867035947740078
step: 80, loss: 0.03860968351364136
step: 90, loss: 0.12528173625469208
step: 100, loss: 0.0304802805185318
step: 110, loss: 0.05744095891714096
step: 120, loss: 0.08530513942241669
step: 130, loss: 0.1975570023059845
step: 140, loss: 0.040452007204294205
step: 150, loss: 0.03381408005952835
step: 160, loss: 0.039303094148635864
step: 170, loss: 0.00640895776450634
step: 180, loss: 0.18249744176864624
```

```
step: 190, loss: 0.05015264451503754
/usr/local/lib/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=0.7835051546391752, f1=0.7570332480818414,
best_f1=0.7570332480818414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.037264734506607056
step: 10, loss: 0.011992444284260273
step: 20, loss: 0.12820032238960266
step: 30, loss: 0.07273861765861511
step: 40, loss: 0.06756807863712311
step: 50, loss: 0.07043226808309555
step: 60, loss: 0.03593083471059799
step: 70, loss: 0.003402020549401641
step: 80, loss: 0.047651588916778564
step: 90, loss: 0.002410756889730692
step: 100, loss: 0.019959915429353714
step: 110, loss: 0.09216293692588806
step: 120, loss: 0.0045075854286551476
step: 130, loss: 0.013031828217208385
step: 140, loss: 0.0041555664502084255
step: 150, loss: 0.025446437299251556
step: 160, loss: 0.004543880000710487
step: 170, loss: 0.007694795727729797
step: 180, loss: 0.12824396789073944
step: 190, loss: 0.15378598868846893
/usr/local/lib/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=0.7955182072829131, f1=0.7477744807121661,
best_f1=0.7477744807121661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01715000346302986
step: 10, loss: 0.000495623389724642
step: 20, loss: 0.038437917828559875
```

```
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 30, loss: 0.0033545419573783875
step: 40, loss: 0.02632709965109825
step: 50, loss: 0.004225711338222027
step: 60, loss: 0.0010908383410423994
step: 70, loss: 0.20870207250118256
step: 80, loss: 0.03159142658114433
step: 90, loss: 0.0582464262843132
step: 100, loss: 0.0213365089148283
step: 110, loss: 0.014209039509296417
step: 120, loss: 0.018432453274726868
step: 130, loss: 0.00648770947009325
step: 140, loss: 0.003274475922808051
step: 150, loss: 0.18815423548221588
step: 160, loss: 0.0015443861484527588
step: 170, loss: 0.019799018278717995
step: 180, loss: 0.011280588805675507
step: 190, loss: 0.014789370819926262
/usr/local/lib/python3.7/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=0.773841961852861, f1=0.7415730337078652,
best_f1=0.7477744807121661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018173910211771727
step: 10, loss: 0.024748530238866806
step: 20, loss: 0.0020641109440475702
step: 30, loss: 0.008925841189920902
step: 40, loss: 0.00956517644226551
step: 50, loss: 0.11989539861679077
step: 60, loss: 0.01357851643115282
step: 70, loss: 0.002310102805495262
step: 80, loss: 0.09827429056167603
step: 90, loss: 0.09405962377786636
step: 100, loss: 0.00496286666020751
step: 110, loss: 0.007255824748426676
step: 120, loss: 0.021318992599844933
step: 130, loss: 0.012167907319962978
step: 140, loss: 0.024494148790836334
step: 150, loss: 0.0829441249370575
step: 160, loss: 0.015015877783298492
step: 170, loss: 0.1808249056339264
step: 180, loss: 0.14386461675167084
step: 190, loss: 0.02195841260254383
/usr/local/lib/python3.7/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=0.7595628415300547, f1=0.732394366197183,
best_f1=0.7477744807121661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010073765879496932
step: 10, loss: 0.011018338613212109
step: 20, loss: 0.0037420305889099836
step: 30, loss: 0.038569312542676926
step: 40, loss: 0.007644126191735268
step: 50, loss: 0.0022336270194500685
step: 60, loss: 0.005612688604742289
step: 70, loss: 0.00091642024926841 26
step: 80, loss: 0.03339245170354843
step: 90, loss: 0.027882713824510574
step: 100, loss: 0.027852308005094528
step: 110, loss: 0.0006925016059540212
step: 120, loss: 0.0036802173126488924
step: 130, loss: 0.0014116086531430483
step: 140, loss: 0.001187330693937838
step: 150, loss: 0.07530529052019119
step: 160, loss: 0.09109488874673843
step: 170, loss: 0.0070081851445138454
step: 180, loss: 0.1282816231250763
step: 190, loss: 0.02578919008374214
/usr/local/lib/python3.7/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=0.7634408602150539, f1=0.7415730337078652,
best_f1=0.7477744807121661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018056515837088227
step: 10, loss: 0.00392201216891408
step: 20, loss: 0.05802619084715843
step: 30, loss: 0.008716368116438389
step: 40, loss: 0.004059193190187216
step: 50, loss: 0.0018223485676571727
step: 60, loss: 0.0034100909251719713
step: 70, loss: 0.004652373492717743
step: 80, loss: 0.0033831270411610603
step: 90, loss: 0.005561815109103918
step: 100, loss: 0.028529874980449677
step: 110, loss: 0.0011802534572780132
step: 120, loss: 0.0015149225946515799
step: 130, loss: 0.0017723587807267904
```

```
step: 140, loss: 0.0006267218268476427
step: 150, loss: 0.0085832504555583
step: 160, loss: 0.0008972353534772992
step: 170, loss: 0.0044037653133273125
step: 180, loss: 0.19263774156570435
step: 190, loss: 0.005198847036808729
/usr/local/lib/python3.7/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=0.7571801566579635, f1=0.7425474254742548,
best_f1=0.7477744807121661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011219073785468936
step: 10, loss: 0.0007967553101480007
step: 20, loss: 0.0020127883180975914
step: 30, loss: 0.003670080564916134
step: 40, loss: 0.0035775909200310707
step: 50, loss: 0.01578625477850437
step: 60, loss: 0.029838109388947487
step: 70, loss: 0.027378754690289497
step: 80, loss: 0.009748273529112339
step: 90, loss: 0.004479188937693834
step: 100, loss: 0.0019798786379396915
step: 110, loss: 0.00576889468356967
step: 120, loss: 0.0013162973336875439
step: 130, loss: 0.11503084003925323
step: 140, loss: 0.013498056679964066
step: 150, loss: 0.059850554913282394
step: 160, loss: 0.010670104995369911
step: 170, loss: 0.017363591119647026
step: 180, loss: 0.006904295179992914
step: 190, loss: 0.007446389645338585
/usr/local/lib/python3.7/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=0.772845953002611, f1=0.7606382978723404,
best_f1=0.7477744807121661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006436080439016223
step: 10, loss: 0.0007299415301531553
step: 20, loss: 0.0022660342510789633
step: 30, loss: 0.007986404933035374
step: 40, loss: 0.0005639577284455299
step: 50, loss: 0.011773496866226196
```

```
step: 60, loss: 0.001417581457644701
step: 70, loss: 0.0003315811918582767
step: 80, loss: 0.011523704044520855
step: 90, loss: 0.007658648770302534
step: 100, loss: 0.0006173709407448769
step: 110, loss: 0.0031072190031409264
step: 120, loss: 0.0016082879155874252
step: 130, loss: 0.00059516960754990058
step: 140, loss: 0.0007688429323025048
step: 150, loss: 0.005525651853531599
step: 160, loss: 0.1153932437300682
step: 170, loss: 0.004758684430271387
step: 180, loss: 0.05112665891647339
step: 190, loss: 0.001578081981278956
/usr/local/lib/python3.7/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=0.7582417582417582, f1=0.7541899441340784,
best_f1=0.7477744807121661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003073714906349778
step: 10, loss: 0.0008033999474719167
step: 20, loss: 0.016523057594895363
step: 30, loss: 0.0034570961724966764
step: 40, loss: 0.003550647757947445
step: 50, loss: 0.00881340354681015
step: 60, loss: 0.00046851104707457125
step: 70, loss: 0.021746492013335228
step: 80, loss: 0.0014976955717429519
step: 90, loss: 0.0016903197392821312
step: 100, loss: 0.000314374832669273
step: 110, loss: 0.0007973459432832897
step: 120, loss: 0.000462940864963457
step: 130, loss: 0.00042044490692205727
step: 140, loss: 0.00569783104583621
step: 150, loss: 0.0009820471750572324
step: 160, loss: 0.002040886553004384
step: 170, loss: 0.002504674019291997
step: 180, loss: 0.0010109114227816463
step: 190, loss: 0.011204730719327927
/usr/local/lib/python3.7/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=0.7596899224806202, f1=0.7506561679790026,
best_f1=0.7477744807121661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006274977931752801
step: 10, loss: 0.0025460312608629465
step: 20, loss: 0.03371996060013771
step: 30, loss: 0.0016685420414432883
step: 40, loss: 0.00026387759135104716
step: 50, loss: 0.04871959984302521
step: 60, loss: 0.0015626748790964484
step: 70, loss: 0.022388245910406113
step: 80, loss: 0.01328798197209835
step: 90, loss: 0.0016245327424257994
step: 100, loss: 0.11933661252260208
step: 110, loss: 0.0006739668315276504
step: 120, loss: 0.009347640909254551
step: 130, loss: 0.0010006553493440151
step: 140, loss: 0.010427206754684448
step: 150, loss: 0.0010045821545645595
step: 160, loss: 0.0007288595661520958
step: 170, loss: 0.000359757017577067
step: 180, loss: 0.0012639124179258943
step: 190, loss: 0.000543635745998472
/usr/local/lib/python3.7/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=0.7602339181286548, f1=0.7426900584795321,
best_f1=0.7477744807121661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.039251234382390976
step: 10, loss: 0.002428448759019375
step: 20, loss: 0.08754298090934753
step: 30, loss: 0.0017060383688658476
step: 40, loss: 0.003534181509166956
step: 50, loss: 0.0034746306482702494
step: 60, loss: 0.0007263047154992819
step: 70, loss: 0.001007901388220489
step: 80, loss: 0.0010955301113426685
step: 90, loss: 0.002646929584443569
step: 100, loss: 0.0003694930928759277
step: 110, loss: 0.001018116483464837
step: 120, loss: 0.0007492281147278845
step: 130, loss: 0.0002975116658490151
step: 140, loss: 0.07339197397232056
step: 150, loss: 0.0006895982660353184
step: 160, loss: 0.0008984742453321815
step: 170, loss: 0.02040843479335308

```
step: 180, loss: 0.007018741685897112
step: 190, loss: 0.001378574874252081
/usr/local/lib/python3.7/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=0.7592067988668556, f1=0.7449856733524355,
best_f1=0.7477744807121661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007983515970408916
step: 10, loss: 0.002298640087246895
step: 20, loss: 0.0014499344397336245
step: 30, loss: 0.0016613121842965484
step: 40, loss: 0.05960717052221298
step: 50, loss: 0.0020764973014593124
step: 60, loss: 0.0034954517614096403
step: 70, loss: 0.0007321984739974141
step: 80, loss: 0.0008079636027105153
step: 90, loss: 0.003759878221899271
step: 100, loss: 0.0008432170143350959
step: 110, loss: 0.0061727771535515785
step: 120, loss: 0.0021314870100468397
step: 130, loss: 0.001190690672956407
step: 140, loss: 0.02230469137430191
step: 150, loss: 0.0007204575231298804
step: 160, loss: 0.0013713750522583723
step: 170, loss: 0.05092189088463783
step: 180, loss: 0.0026957057416439056
step: 190, loss: 0.00043660460505634546
/usr/local/lib/python3.7/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=0.76, f1=0.7341040462427746, best_f1=0.7477744807121661
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.")
2049it [00:07, 279.84it/s]
load_f1 = 0.7319587628865978
real_f1 = 0.7125307125307127
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 [00:12, 349.24it/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 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: 0.8516586422920227
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.2231062948703766
step: 20, loss: 0.1499432921409607
step: 30, loss: 0.24753202497959137
step: 40, loss: 0.31307488679885864
step: 50, loss: 0.3790012300014496
step: 60, loss: 0.44139033555984497
step: 70, loss: 0.31067919731140137
step: 80, loss: 0.25100111961364746
step: 90, loss: 0.40560510754585266
step: 100, loss: 0.2283898890018463
step: 110, loss: 0.1915871948003769
step: 120, loss: 0.5493771433830261
step: 130, loss: 0.4236079156398773
step: 140, loss: 0.48431774973869324
step: 150, loss: 0.1122460812330246
step: 160, loss: 0.33719369769096375
step: 170, loss: 0.2119043916463852
/usr/local/lib/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=0.5529715762273902, f1=0.544502617801047,
best_f1=0.544502617801047
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31814807653427124
step: 10, loss: 0.15517506003379822
step: 20, loss: 0.28977248072624207
step: 30, loss: 0.24094879627227783
step: 40, loss: 0.1838684380054474
step: 50, loss: 0.1831911951303482
step: 60, loss: 0.07428029924631119
step: 70, loss: 0.27987897396087646
step: 80, loss: 0.052668508142232895
step: 90, loss: 0.19681592285633087
step: 100, loss: 0.21974216401576996
step: 110, loss: 0.2666928768157959

```
step: 120, loss: 0.06693783402442932
step: 130, loss: 0.06440769881010056
step: 140, loss: 0.14428110420703888
step: 150, loss: 0.14818617701530457
step: 160, loss: 0.12295973300933838
step: 170, loss: 0.3168042302131653
/usr/local/lib/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=0.6929133858267718, f1=0.7241379310344829,
best_f1=0.7241379310344829
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17044641077518463
step: 10, loss: 0.08774730563163757
step: 20, loss: 0.04855648800730705
step: 30, loss: 0.4056174159049988
step: 40, loss: 0.025459127500653267
step: 50, loss: 0.07258298248052597
step: 60, loss: 0.28759604692459106
step: 70, loss: 0.12162267416715622
step: 80, loss: 0.3251765966415405
step: 90, loss: 0.0735739916563034
step: 100, loss: 0.02504381351172924
step: 110, loss: 0.08753670752048492
step: 120, loss: 0.04681963846087456
step: 130, loss: 0.1633889377117157
step: 140, loss: 0.023422980681061745
step: 150, loss: 0.08480074256658554
step: 160, loss: 0.06186869740486145
step: 170, loss: 0.1532488465309143
/usr/local/lib/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=0.7310344827586207, f1=0.7168949771689498,
best_f1=0.7168949771689498
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.06058177351951599
step: 10, loss: 0.13059109449386597
step: 20, loss: 0.16683319211006165
step: 30, loss: 0.0727161094546318
step: 40, loss: 0.2558300197124481
step: 50, loss: 0.020692922174930573
step: 60, loss: 0.09915145486593246
step: 70, loss: 0.023013243451714516
step: 80, loss: 0.03397376462817192
step: 90, loss: 0.029145434498786926
step: 100, loss: 0.013797915540635586
step: 110, loss: 0.04354812949895859
step: 120, loss: 0.19062387943267822
step: 130, loss: 0.05603644251823425
step: 140, loss: 0.026111997663974762
step: 150, loss: 0.023160332813858986
step: 160, loss: 0.11976661533117294
step: 170, loss: 0.06761102378368378
/usr/local/lib/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=0.7621247113163971, f1=0.735042735042735,
best_f1=0.735042735042735
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06839722394943237
step: 10, loss: 0.0866074487566948
step: 20, loss: 0.02374856173992157
step: 30, loss: 0.10162952542304993
step: 40, loss: 0.006142972037196159
step: 50, loss: 0.062867671251297
step: 60, loss: 0.009262342937290668
step: 70, loss: 0.03807108849287033
step: 80, loss: 0.049853336066007614
step: 90, loss: 0.01034468598663807
step: 100, loss: 0.023676740005612373
step: 110, loss: 0.15099938213825226
step: 120, loss: 0.029555002227425575
step: 130, loss: 0.02101709134876728
step: 140, loss: 0.060692548751831055
step: 150, loss: 0.017793942242860794
step: 160, loss: 0.02091178670525551
step: 170, loss: 0.09417688846588135
```

```
/usr/local/lib/python3.7/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=0.7218045112781954, f1=0.755868544600939,
best_f1=0.735042735042735
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025744110345840454
step: 10, loss: 0.037519246339797974
step: 20, loss: 0.010879194363951683
step: 30, loss: 0.09027412533760071
step: 40, loss: 0.005511933472007513
step: 50, loss: 0.04489472508430481
step: 60, loss: 0.007458607666194439
step: 70, loss: 0.023466316983103752
step: 80, loss: 0.16817010939121246
step: 90, loss: 0.010805279016494751
step: 100, loss: 0.00492337578907609
step: 110, loss: 0.013236945495009422
step: 120, loss: 0.08439166843891144
step: 130, loss: 0.24798165261745453
step: 140, loss: 0.02661127597093582
step: 150, loss: 0.15803250670433044
step: 160, loss: 0.01573093608021736
step: 170, loss: 0.002837425097823143
/usr/local/lib/python3.7/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=0.7382550335570469, f1=0.7234042553191489,
best_f1=0.735042735042735
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002937889890745282
step: 10, loss: 0.0020515071228146553
step: 20, loss: 0.008990665897727013
step: 30, loss: 0.0027132006362080574
step: 40, loss: 0.0036265612579882145
step: 50, loss: 0.008528000675141811
step: 60, loss: 0.16300693154335022
step: 70, loss: 0.004240274894982576
step: 80, loss: 0.0009824603330343962
step: 90, loss: 0.013150861486792564
step: 100, loss: 0.00715660722926259
step: 110, loss: 0.0025488920509815216
step: 120, loss: 0.10701858997344971
step: 130, loss: 0.5455176830291748
```

```
step: 140, loss: 0.021655051037669182
step: 150, loss: 0.09727510064840317
step: 160, loss: 0.023380350321531296
step: 170, loss: 0.18974313139915466
/usr/local/lib/python3.7/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=0.7400468384074942, f1=0.7346938775510203,
best_f1=0.735042735042735
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013415698893368244
step: 10, loss: 0.05145612731575966
step: 20, loss: 0.016148194670677185
step: 30, loss: 0.02667658030986786
step: 40, loss: 0.001966835232451558
step: 50, loss: 0.011010355316102505
step: 60, loss: 0.006831732578575611
step: 70, loss: 0.043739087879657745
step: 80, loss: 0.003311433829367161
step: 90, loss: 0.03609530255198479
step: 100, loss: 0.0014708245871588588
step: 110, loss: 0.021618777886033058
step: 120, loss: 0.012737280689179897
step: 130, loss: 0.0035237933043390512
step: 140, loss: 0.004545470699667931
step: 150, loss: 0.10939241945743561
step: 160, loss: 0.0890023410320282
step: 170, loss: 0.0035471394658088684
/usr/local/lib/python3.7/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=0.7604395604395604, f1=0.7569892473118278,
best_f1=0.735042735042735
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017492331564426422
step: 10, loss: 0.020693611353635788
step: 20, loss: 0.02425961382687092
step: 30, loss: 0.039227940142154694
step: 40, loss: 0.030529353767633438
step: 50, loss: 0.033150676637887955
step: 60, loss: 0.08216656744480133
step: 70, loss: 0.03115459531545639
step: 80, loss: 0.05527126416563988
step: 90, loss: 0.004929161164909601
```

step: 100, loss: 0.015110401436686516
step: 110, loss: 0.03817898407578468
step: 120, loss: 0.004273451864719391
step: 130, loss: 0.0009478880674578249
step: 140, loss: 0.0012596158776432276
step: 150, loss: 0.02793586254119873
step: 160, loss: 0.09367068111896515
step: 170, loss: 0.018845656886696815
/usr/local/lib/python3.7/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=0.7386934673366835, f1=0.7391304347826086,
best_f1=0.735042735042735
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03261378034949303
step: 10, loss: 0.03323933109641075
step: 20, loss: 0.0023057840298861265
step: 30, loss: 0.012266261503100395
step: 40, loss: 0.013742683455348015
step: 50, loss: 0.08407115936279297
step: 60, loss: 0.005286895204335451
step: 70, loss: 0.16725395619869232
step: 80, loss: 0.024665242061018944
step: 90, loss: 0.016594097018241882
step: 100, loss: 0.06215474754571915
step: 110, loss: 0.030182678252458572
step: 120, loss: 0.03560882434248924
step: 130, loss: 0.038446925580501556
step: 140, loss: 0.02873239666223526
step: 150, loss: 0.02826986089348793
step: 160, loss: 0.0031437843572348356
step: 170, loss: 0.029078569263219833
/usr/local/lib/python3.7/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=0.730077120822622, f1=0.7677261613691931,
best_f1=0.735042735042735
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003738616651389748
step: 10, loss: 0.005499484017491341
step: 20, loss: 0.025279266759753227
step: 30, loss: 0.0014377456391230226
step: 40, loss: 0.018006999045610428
step: 50, loss: 0.002944991225376725

```
step: 60, loss: 0.028883587568998337
step: 70, loss: 0.00579936895519495
step: 80, loss: 0.0029942234978079796
step: 90, loss: 0.0019159232033416629
step: 100, loss: 0.00528641976416111
step: 110, loss: 0.014474782161414623
step: 120, loss: 0.04385926201939583
step: 130, loss: 0.0037552148569375277
step: 140, loss: 0.06850270926952362
step: 150, loss: 0.008523895405232906
step: 160, loss: 0.0026333665009588003
step: 170, loss: 0.08859139680862427
/usr/local/lib/python3.7/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=0.7254408060453401, f1=0.7662650602409637,
best_f1=0.735042735042735
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008882632479071617
step: 10, loss: 0.0028059757314622402
step: 20, loss: 0.006547632627189159
step: 30, loss: 0.001801114296540618
step: 40, loss: 0.0002783258387353271
step: 50, loss: 0.0024964138865470886
step: 60, loss: 0.0009981192415580153
step: 70, loss: 0.013485400937497616
step: 80, loss: 0.025680813938379288
step: 90, loss: 0.0004382602055557072
step: 100, loss: 0.0007068405975587666
step: 110, loss: 0.000276327773462981
step: 120, loss: 0.025759795680642128
step: 130, loss: 0.016224097460508347
step: 140, loss: 0.0005778901977464557
step: 150, loss: 0.043979983776807785
step: 160, loss: 0.013723691925406456
step: 170, loss: 0.0006508453516289592
/usr/local/lib/python3.7/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=0.7365853658536585, f1=0.7610208816705337,
best_f1=0.735042735042735
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015087929787114263
step: 10, loss: 0.2563602924346924
```

```
step: 20, loss: 0.001035527791827917
step: 30, loss: 0.007387390825897455
step: 40, loss: 0.002723419340327382
step: 50, loss: 0.002178365597501397
step: 60, loss: 0.008151291869580746
step: 70, loss: 0.014688648283481598
step: 80, loss: 0.0038914652541279793
step: 90, loss: 0.06957570463418896
step: 100, loss: 0.05382377654314041
step: 110, loss: 0.03206607326865196
step: 120, loss: 0.002270917873829603
step: 130, loss: 0.006030393298715353
step: 140, loss: 0.0017971681663766503
step: 150, loss: 0.0004276365216355771
step: 160, loss: 0.005932836793363094
step: 170, loss: 0.0015801741974428296
/usr/local/lib/python3.7/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=0.7360774818401936, f1=0.7505720823798627,
best_f1=0.735042735042735
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0045503126457333565
step: 10, loss: 0.00038990742177702487
step: 20, loss: 0.006083348300307989
step: 30, loss: 0.0011514727957554671
step: 40, loss: 0.027318106964230537
step: 50, loss: 0.014810390770435333
step: 60, loss: 0.00041786272777724445
step: 70, loss: 0.009239207953214645
step: 80, loss: 0.004753818269819021
step: 90, loss: 0.0010862428462132812
step: 100, loss: 0.0032262816093862057
step: 110, loss: 0.026194149628281593
step: 120, loss: 0.0038603341672569513
step: 130, loss: 0.0015985050704330206
step: 140, loss: 0.0011370478896424174
step: 150, loss: 0.0009306185529567301
step: 160, loss: 0.03341438248753548
step: 170, loss: 0.00038829119876027107
/usr/local/lib/python3.7/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=0.7321867321867321, f1=0.7546296296296297,
best_f1=0.735042735042735
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0054680174216628075
step: 10, loss: 0.004949726164340973
step: 20, loss: 0.039564840495586395
step: 30, loss: 0.010455445386469364
step: 40, loss: 0.000226815405767411
step: 50, loss: 0.0019919329788535833
step: 60, loss: 0.0002655299613252282
step: 70, loss: 0.009757587686181068
step: 80, loss: 0.0021992248948663473
step: 90, loss: 0.00017956807278096676
step: 100, loss: 0.0005942534189671278
step: 110, loss: 0.0005129189230501652
step: 120, loss: 0.006895285565406084
step: 130, loss: 0.0015088655054569244
step: 140, loss: 0.007626861333847046
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 150, loss: 0.005954256746917963
step: 160, loss: 0.0004662301216740161
step: 170, loss: 0.016529595479369164
/usr/local/lib/python3.7/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=0.7350835322195706, f1=0.7505720823798627,
best_f1=0.735042735042735
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.")
1916it [00:05, 337.31it/s]
load_f1 = 0.5555555555555556
real_f1 = 0.5272331154684097
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, 259.76it/s]
```

4.3   DITTO DIRTY

4.3.1   DBLP-ACM - Running the matcher

```python
[ ]:  # 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 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: 0.7975496649742126
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4776090979576111

```
step: 20, loss: 0.6129978895187378
step: 30, loss: 0.48991456627845764
step: 40, loss: 0.3799011707305908
step: 50, loss: 0.23010683059692383
step: 60, loss: 0.1712636947631836
step: 70, loss: 0.12422790378332138
step: 80, loss: 0.3063082993030548
step: 90, loss: 0.15512147545814514
step: 100, loss: 0.1774204820394516
step: 110, loss: 0.15438073873519897
step: 120, loss: 0.0696462020277977
step: 130, loss: 0.004681678488850594
step: 140, loss: 0.08342385292053223
step: 150, loss: 0.12719139456748962
step: 160, loss: 0.22824802994728088
step: 170, loss: 0.014853611588478088
step: 180, loss: 0.011707672849297523
step: 190, loss: 0.2114706039428711
step: 200, loss: 0.01720821112394333
step: 210, loss: 0.022739792242646217
step: 220, loss: 0.007983185350894928
step: 230, loss: 0.0062584057450294495
/usr/local/lib/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=0.9514348785871964, f1=0.9540873460246361,
best_f1=0.9540873460246361
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.175237059593200068
step: 10, loss: 0.07619182020425797
step: 20, loss: 0.12075842171907425
step: 30, loss: 0.03064686432480812
step: 40, loss: 0.020272448658943176
step: 50, loss: 0.028645765036344528
step: 60, loss: 0.10178423672914505
step: 70, loss: 0.10641543567180634
step: 80, loss: 0.017038755118846893
step: 90, loss: 0.0033488355111330748
step: 100, loss: 0.039643656462430954
step: 110, loss: 0.03359577804803848
step: 120, loss: 0.06126171350479126
step: 130, loss: 0.0064893607050180435
```

```
step: 140, loss: 0.056398604065179825
step: 150, loss: 0.00587540864944458
step: 160, loss: 0.020391544327139854
step: 170, loss: 0.02182321809232235
step: 180, loss: 0.02517147921025753
step: 190, loss: 0.03312380611896515
step: 200, loss: 0.17238202691078186
step: 210, loss: 0.11279658228158951
step: 220, loss: 0.0060882833786308765
step: 230, loss: 0.0057729980908334255
/usr/local/lib/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=0.961625282167043, f1=0.9683972911963882,
best_f1=0.9683972911963882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.036041509360075
step: 10, loss: 0.04198283702135086
step: 20, loss: 0.019797977060079575
step: 30, loss: 0.060534797608852386
step: 40, loss: 0.1318727284669876
step: 50, loss: 0.003268466331064701
step: 60, loss: 0.2375408411026001
step: 70, loss: 0.0041931201703846455
step: 80, loss: 0.10316916555166245
step: 90, loss: 0.04499863088130951
step: 100, loss: 0.003320159390568733
step: 110, loss: 0.05680806189775467
step: 120, loss: 0.015267216600477695
step: 130, loss: 0.005900819785892963
step: 140, loss: 0.0021074258256703615
step: 150, loss: 0.0020085840951651335
step: 160, loss: 0.008146421983838081
step: 170, loss: 0.057570863515138626
step: 180, loss: 0.0019349748035892844
step: 190, loss: 0.0345148928463459
step: 200, loss: 0.007284167688339949
step: 210, loss: 0.0013892694842070341
step: 220, loss: 0.003996757324784994
step: 230, loss: 0.22723926603794098
/usr/local/lib/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=0.9669327251995439, f1=0.9681093394077448,
best_f1=0.9681093394077448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003786563640460372
step: 10, loss: 0.008454159833490849
step: 20, loss: 0.031276535242795944
step: 30, loss: 0.004818392917513847
step: 40, loss: 0.08115024119615555
step: 50, loss: 0.0043100821785628796
step: 60, loss: 0.0015844452427700162
step: 70, loss: 0.0921163335442543
step: 80, loss: 0.20214977860450745
step: 90, loss: 0.014032193459570408
step: 100, loss: 0.0017260677414014935
step: 110, loss: 0.06651058793067932
step: 120, loss: 0.07281100004911423
step: 130, loss: 0.0501159243285656
step: 140, loss: 0.010920455679297447
step: 150, loss: 0.010710607282817364
step: 160, loss: 0.001721858629025519
step: 170, loss: 0.004591194447129965
step: 180, loss: 0.02908473275601864
step: 190, loss: 0.031682245433330536
step: 200, loss: 0.02146223932504654
step: 210, loss: 0.007316313683986664
step: 220, loss: 0.0030993553809821606
step: 230, loss: 0.000938547367695719
/usr/local/lib/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=0.9688195991091313, f1=0.9710467706013363,
best_f1=0.9710467706013363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007773788529448211
step: 10, loss: 0.0011493233032524586
```

```
step: 20, loss: 0.0027741435915231705
step: 30, loss: 0.0005961976712569594
step: 40, loss: 0.0015367163578048348
step: 50, loss: 0.0005225921631790698
step: 60, loss: 0.0008111027418635786
step: 70, loss: 0.0017790062120184302
step: 80, loss: 0.002157966373488307
step: 90, loss: 0.0038570372853428125
step: 100, loss: 0.011726095341145992
step: 110, loss: 0.000941043661441654
step: 120, loss: 0.03687281906604767
step: 130, loss: 0.03365505486726761
step: 140, loss: 0.004256019834429026
step: 150, loss: 0.0005372338928282261
step: 160, loss: 0.056173175573349
step: 170, loss: 0.0592663437128067
step: 180, loss: 0.00392730999737978
step: 190, loss: 0.0015573021955788136
step: 200, loss: 0.0044418396428227425
step: 210, loss: 0.013991836458444595
step: 220, loss: 0.001081372145563364
step: 230, loss: 0.02838345617055893
/usr/local/lib/python3.7/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=0.9666666666666666, f1=0.9721913236929923,
best_f1=0.9710467706013363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01021286379545927
step: 10, loss: 0.0053793080151081085
step: 20, loss: 0.0004548681026790291
step: 30, loss: 0.003229917958378792
step: 40, loss: 0.006178226321935654
step: 50, loss: 0.04104554280638695
step: 60, loss: 0.0013595663476735353
step: 70, loss: 0.018408682197332382
step: 80, loss: 0.0010231632040813565
step: 90, loss: 0.00023556123778689653
step: 100, loss: 0.00016438662714790553
step: 110, loss: 0.0007253499352373183
step: 120, loss: 0.0007297255215235054
step: 130, loss: 0.003015427850186825
step: 140, loss: 0.0005317747127264738
step: 150, loss: 0.0014369722921401262
step: 160, loss: 0.002927024383097887
step: 170, loss: 0.013119040988385677
```

```
step: 180, loss: 0.0010034095030277967
step: 190, loss: 0.00607543857768178
step: 200, loss: 0.09068869799375534
step: 210, loss: 0.0008846729178912938
step: 220, loss: 0.00023508943559136242
step: 230, loss: 0.0004313874524086714
/usr/local/lib/python3.7/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=0.9666666666666666, f1=0.9700332963374029,
best_f1=0.9710467706013363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.036312539130449295
step: 10, loss: 0.0008780673379078507
step: 20, loss: 0.002440913813188672
step: 30, loss: 0.0013889140682294965
step: 40, loss: 0.052334200059084892
step: 50, loss: 0.0023170909844338894
step: 60, loss: 0.020983733236789703
step: 70, loss: 0.0030306673143059015
step: 80, loss: 0.0005830627051182091
step: 90, loss: 0.0003496795834507793
step: 100, loss: 0.0003112567064817995
step: 110, loss: 0.0009315941715613008
step: 120, loss: 0.056055061519145966
step: 130, loss: 0.00038306351052597165
step: 140, loss: 0.00033853319473564625
step: 150, loss: 0.0010062640067189932
step: 160, loss: 0.00023000320652499795
step: 170, loss: 0.0023720632307231426
step: 180, loss: 0.017940703779459
step: 190, loss: 0.002586898161098361
step: 200, loss: 0.000841617351397872
step: 210, loss: 0.000182329211384058
step: 220, loss: 0.0014446023851633072
step: 230, loss: 0.033461470156908035
/usr/local/lib/python3.7/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=0.9578713968957872, f1=0.9688195991091313,
best_f1=0.9710467706013363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013381954282522202
step: 10, loss: 0.004693920258432627
```

```
step: 20, loss: 0.00048162543680518866
step: 30, loss: 0.00022395525593310595
step: 40, loss: 0.0029062852263450623
step: 50, loss: 0.0003300231765024364
step: 60, loss: 0.005885547958314419
step: 70, loss: 0.000244892988121137
step: 80, loss: 0.0003524889179971069
step: 90, loss: 0.010718903504312038
step: 100, loss: 0.0001237285468960181
step: 110, loss: 0.00014186288171913475
step: 120, loss: 0.0009406860917806625
step: 130, loss: 0.0025150130968540907
step: 140, loss: 0.0008014337508939207
step: 150, loss: 0.0003080359019804746
step: 160, loss: 0.005951344035565853
step: 170, loss: 0.00026004138635471463
step: 180, loss: 0.007977907545864582
step: 190, loss: 0.0018696993356570601
step: 200, loss: 0.02758997492492199
step: 210, loss: 0.0002535871753934771
step: 220, loss: 0.0003150130796711892
step: 230, loss: 0.02718494087457657
/usr/local/lib/python3.7/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=0.9623059866962307, f1=0.9645232815964524,
best_f1=0.9710467706013363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002529750927351415
step: 10, loss: 0.014209264889359474
step: 20, loss: 0.0011792326113209128
step: 30, loss: 0.00023796963796485215
step: 40, loss: 8.094509394140914e-05
step: 50, loss: 0.0014445193810388446
step: 60, loss: 0.0005640155868604779
step: 70, loss: 0.0017478910740464926
step: 80, loss: 0.0004766362835653126
step: 90, loss: 0.000382458878448233
step: 100, loss: 0.00792563334107399
step: 110, loss: 8.748685650061816e-05
step: 120, loss: 0.07835837453603745
step: 130, loss: 0.0004025857779197395
step: 140, loss: 0.004329704213887453
step: 150, loss: 5.9054087614640594e-05
step: 160, loss: 0.0008109157788567245
step: 170, loss: 7.03239202266559e-05
```

```
step: 180, loss: 0.002390407957136631
step: 190, loss: 0.0001563989499118179
step: 200, loss: 9.991783008445054e-05
step: 210, loss: 0.00040137459291145205
step: 220, loss: 0.002877661259844899
step: 230, loss: 0.00048641767352819443
/usr/local/lib/python3.7/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=0.9587513935340021, f1=0.9710467706013363,
best_f1=0.9710467706013363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013197892985772341
step: 10, loss: 9.872922237263992e-05
step: 20, loss: 0.0003795016964431852
step: 30, loss: 0.00018832759815268219
step: 40, loss: 0.0012993670534342527
step: 50, loss: 0.00027071655495092273
step: 60, loss: 0.009386168792843819
step: 70, loss: 0.0005224659689702094
step: 80, loss: 0.00045512878568843007
step: 90, loss: 0.00022014993010088801
step: 100, loss: 6.77838470437564e-05
step: 110, loss: 0.00013745152682531625
step: 120, loss: 0.01716359704732895
step: 130, loss: 0.002050766721367836
step: 140, loss: 0.0013092380249872804
step: 150, loss: 0.023182732984423637
step: 160, loss: 0.0008828595746308565
step: 170, loss: 0.00020544846483971924
step: 180, loss: 0.05851467326283455
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 190, loss: 0.00030129082733765244
step: 200, loss: 0.00016923325893003494
step: 210, loss: 0.005294067785143852
step: 220, loss: 0.0016875392757356167
step: 230, loss: 0.000579710234887898
/usr/local/lib/python3.7/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=0.9608062709966406, f1=0.9652855543113102,
best_f1=0.9710467706013363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006917256396263838
```

```
step: 10, loss: 0.000697015319019556
step: 20, loss: 8.800127397989854e-05
step: 30, loss: 0.0009701072704046965
step: 40, loss: 0.021225186064839363
step: 50, loss: 0.006035068538039923
step: 60, loss: 0.00036262592766433954
step: 70, loss: 0.0005612067761830986
step: 80, loss: 0.0001617989910300821
step: 90, loss: 0.0015983706107363105
step: 100, loss: 0.0003004073223564774
step: 110, loss: 0.0007428441895172
step: 120, loss: 0.007877555675804615
step: 130, loss: 6.413116352632642e-05
step: 140, loss: 0.0010596953798085451
step: 150, loss: 0.00012600322952494025
step: 160, loss: 0.0019581844098865986
step: 170, loss: 0.0025712724309414625
step: 180, loss: 0.002264623995870352
step: 190, loss: 0.00031212467001751065
step: 200, loss: 0.0004609323223121464
step: 210, loss: 9.744679118739441e-05
step: 220, loss: 0.0008216226124204695
step: 230, loss: 0.0013968496350571513
/usr/local/lib/python3.7/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=0.9607182940516273, f1=0.9641255605381166,
best_f1=0.9710467706013363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008305825758725405
step: 10, loss: 0.026585610583424568
step: 20, loss: 0.00010277818364556879
step: 30, loss: 0.001371918711811304
step: 40, loss: 0.00011282571358606219
step: 50, loss: 0.00011807812552433461
step: 60, loss: 0.0002958890690933913
step: 70, loss: 0.0001606019795872271
step: 80, loss: 0.0005796549376100302
step: 90, loss: 3.725198985193856e-05
step: 100, loss: 0.00018741990788839757
step: 110, loss: 0.13174811005592346
step: 120, loss: 0.00013790138473268598
step: 130, loss: 0.0020194421522319317
step: 140, loss: 4.872055069427006e-05
step: 150, loss: 0.0003041085146833211
step: 160, loss: 0.012152711860835552
```

```
step: 170, loss: 0.00019155755580868572
step: 180, loss: 0.00021083744650240988
step: 190, loss: 0.00020853539172094315
step: 200, loss: 0.00019413868722040206
step: 210, loss: 3.201823346898891e-05
step: 220, loss: 0.02605750598013401
step: 230, loss: 0.0002571757941041142
/usr/local/lib/python3.7/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=0.9612403100775194, f1=0.9614112458654906,
best_f1=0.9710467706013363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.498118015471846e-05
step: 10, loss: 0.000233012848184444
step: 20, loss: 0.000130821208585985
step: 30, loss: 0.00010327730706194416
step: 40, loss: 0.00036228442331776023
step: 50, loss: 0.00013743418094236404
step: 60, loss: 0.0013683909783139825
step: 70, loss: 8.680592873133719e-05
step: 80, loss: 0.00028439657762646675
step: 90, loss: 7.072585867717862e-05
step: 100, loss: 8.847707795212045e-05
step: 110, loss: 7.085687684593722e-05
step: 120, loss: 5.8136476582149044e-05
step: 130, loss: 0.000348057656083256
step: 140, loss: 0.0033705325331538916
step: 150, loss: 8.561746653867885e-05
step: 160, loss: 0.0003727259172592312
step: 170, loss: 0.00020511311595328152
step: 180, loss: 0.0045897094532847404
step: 190, loss: 0.0024409021716564894
step: 200, loss: 0.0002386963606113568
step: 210, loss: 6.049950388842262e-05
step: 220, loss: 0.0005074942600913346
step: 230, loss: 7.56301888031885e-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=0.9631284916201117, f1=0.9720670391061451,
best_f1=0.9710467706013363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00033655649167485535
```

```
step: 10, loss: 8.49856369313784e-05
step: 20, loss: 0.004266953095793724
step: 30, loss: 0.00039292036672122777
step: 40, loss: 4.015645390609279e-05
step: 50, loss: 0.0005753469304181635
step: 60, loss: 0.0008089513285085559
step: 70, loss: 0.0001484170788899064
step: 80, loss: 0.0011252093827351928
step: 90, loss: 0.00011392321903258562
step: 100, loss: 0.00015321376849897206
step: 110, loss: 0.005206014961004257
step: 120, loss: 0.00010409914830233902
step: 130, loss: 4.944395914208144e-05
step: 140, loss: 0.0002939913247246295
step: 150, loss: 9.275579213863239e-05
step: 160, loss: 0.00011154027743032202
step: 170, loss: 0.00013453789870254695
step: 180, loss: 0.00026370974956080232
step: 190, loss: 6.369381299009547e-05
step: 200, loss: 7.163842383306473e-05
step: 210, loss: 0.0003788548056036234
step: 220, loss: 6.248150748433545e-05
step: 230, loss: 0.006493803113698959
/usr/local/lib/python3.7/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=0.9632107023411371, f1=0.9699666295884317,
best_f1=0.9710467706013363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012609496479853988
step: 10, loss: 0.00031352919177152216
step: 20, loss: 3.668208955887122e-05
step: 30, loss: 0.0008144896710291505
step: 40, loss: 5.5526168580399826e-05
step: 50, loss: 8.353058365173638e-05
step: 60, loss: 6.855509855085984e-05
step: 70, loss: 5.406943091657013e-05
step: 80, loss: 0.0006225531687960029
step: 90, loss: 4.654465737985447e-05
step: 100, loss: 0.0001051468207151629
step: 110, loss: 0.0074302139692008495
step: 120, loss: 6.29160858807154e-05
step: 130, loss: 0.00017808166739996523
step: 140, loss: 0.017828533425927162
step: 150, loss: 0.006200631149113178
step: 160, loss: 0.001533123548142612
```

```
step: 170, loss: 7.062700024107471e-05
step: 180, loss: 5.55613987671677e-05
step: 190, loss: 0.022055864334106445
step: 200, loss: 8.905782306101173e-05
step: 210, loss: 0.000771828053984791
step: 220, loss: 0.0076834168285131454
step: 230, loss: 0.00040514906868338585
/usr/local/lib/python3.7/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=0.9642857142857144, f1=0.9688195991091313,
best_f1=0.9710467706013363
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 [00:10, 240.69it/s]
load_f1 = 0.9699666295884317
real_f1 = 0.9688195991091313
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, 263.80it/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 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
```

442

```
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: 0.7898697853088379
step: 10, loss: 0.42180803418159485
step: 20, loss: 0.5062073469161987
step: 30, loss: 0.4483245015144348
step: 40, loss: 0.40668222308158875
step: 50, loss: 0.2748856544494629
step: 60, loss: 0.3390170931816101
step: 70, loss: 0.2455739974975586
step: 80, loss: 0.15938504040241241
step: 90, loss: 0.18008236587047577
step: 100, loss: 0.2540983557701111
step: 110, loss: 0.20166020095348358
step: 120, loss: 0.1078135073184967
step: 130, loss: 0.0466604121029377
step: 140, loss: 0.32659560441970825
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 150, loss: 0.033628467470407486
step: 160, loss: 0.22589188814163208
step: 170, loss: 0.2884162962436676
step: 180, loss: 0.11859321594238281

```
step: 190, loss: 0.09323347359895706
step: 200, loss: 0.10328009724617004
step: 210, loss: 0.10761890560388565
step: 220, loss: 0.2629014849662781
step: 230, loss: 0.13411164283752441
step: 240, loss: 0.20070892572402954
step: 250, loss: 0.036507539451122284
step: 260, loss: 0.058341577649116516
step: 270, loss: 0.02906423807144165
step: 280, loss: 0.07706587761640549
step: 290, loss: 0.09621164947748184
step: 300, loss: 0.1941935271024704
step: 310, loss: 0.04680832847952843
step: 320, loss: 0.1469716727733612
step: 330, loss: 0.18081499636173248
step: 340, loss: 0.24586664140224457
step: 350, loss: 0.12218866497278214
step: 360, loss: 0.08082196861505508
step: 370, loss: 0.14101111888885498
step: 380, loss: 0.19973129034042358
step: 390, loss: 0.015755431726574898
step: 400, loss: 0.03416125848889351
step: 410, loss: 0.10051465779542923
step: 420, loss: 0.03631677106022835
step: 430, loss: 0.11903692036867142
step: 440, loss: 0.13474373519420624
step: 450, loss: 0.028868164867162704
step: 460, loss: 0.026543840765953064
step: 470, loss: 0.16582465171813965
step: 480, loss: 0.22961939871311188
step: 490, loss: 0.05834181234240532
step: 500, loss: 0.027237020432949066
step: 510, loss: 0.08094140142202377
step: 520, loss: 0.04757634922862053
step: 530, loss: 0.06589652597904205
/usr/local/lib/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=0.9170263788968827, f1=0.9113323850165956,
best_f1=0.9113323850165956
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1206289753317833
```

```
step: 10, loss: 0.09123554080724716
step: 20, loss: 0.0936218649148941
step: 30, loss: 0.09111884981393814
step: 40, loss: 0.013462780974805355
step: 50, loss: 0.01348071824759245
step: 60, loss: 0.3322310149669647
step: 70, loss: 0.10785743594169617
step: 80, loss: 0.02483939565718174
step: 90, loss: 0.04274832084774971
step: 100, loss: 0.2169356346130371
step: 110, loss: 0.02360564097762108
step: 120, loss: 0.053847212344408035
step: 130, loss: 0.044439010322093964
step: 140, loss: 0.1119125634431839
step: 150, loss: 0.04794861003756523
step: 160, loss: 0.027867237105965614
step: 170, loss: 0.09403835982084274
step: 180, loss: 0.047532856464385986
step: 190, loss: 0.044891562312841415
step: 200, loss: 0.15861395001411438
step: 210, loss: 0.07301005721092224
step: 220, loss: 0.1137990951538086
step: 230, loss: 0.013010860420763493
step: 240, loss: 0.08753306418657303
step: 250, loss: 0.18007422983646393
step: 260, loss: 0.020184965804219246
step: 270, loss: 0.15293577313423157
step: 280, loss: 0.12253066897392273
step: 290, loss: 0.06398728489875793
step: 300, loss: 0.06622343510389328
step: 310, loss: 0.052800580859184265
step: 320, loss: 0.15390422940254211
step: 330, loss: 0.028602786362171173
step: 340, loss: 0.00798124261200428
step: 350, loss: 0.08698638528585434
step: 360, loss: 0.23641064763069153
step: 370, loss: 0.016577482223510742
step: 380, loss: 0.11585384607315063
step: 390, loss: 0.038672495633363724
step: 400, loss: 0.0653691366314888
step: 410, loss: 0.0017705330392345786
step: 420, loss: 0.062348220497369766
step: 430, loss: 0.01664949208498001
step: 440, loss: 0.022522609680891037
step: 450, loss: 0.03161558881402016
step: 460, loss: 0.24176853895187378
step: 470, loss: 0.021976258605718613
step: 480, loss: 0.12746861577033997
```

```
step: 490, loss: 0.1252015382051468
step: 500, loss: 0.03077445738017559
step: 510, loss: 0.09946313500404358
step: 520, loss: 0.11609674990177155
step: 530, loss: 0.29853421449661255
/usr/local/lib/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=0.9233593391463975, f1=0.9190172884440401,
best_f1=0.9190172884440401
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024387352168560028
step: 10, loss: 0.02545301802456379
step: 20, loss: 0.2773590385913849
step: 30, loss: 0.16412264108657837
step: 40, loss: 0.00642075389623642
step: 50, loss: 0.08629783987998962
step: 60, loss: 0.03515513613820076
step: 70, loss: 0.05611146613955498
step: 80, loss: 0.01561796199530363
step: 90, loss: 0.07784485816955566
step: 100, loss: 0.021753260865807533
step: 110, loss: 0.01241523027420044
step: 120, loss: 0.05896414816379547
step: 130, loss: 0.015265267342329025
step: 140, loss: 0.06722313910722733
step: 150, loss: 0.005547086708247662
step: 160, loss: 0.0034412145614624023
step: 170, loss: 0.03745901212096214
step: 180, loss: 0.006705513224005699
step: 190, loss: 0.0068525332026183605
step: 200, loss: 0.09675902128219604
step: 210, loss: 0.06659495085477829
step: 220, loss: 0.010278399102389812
step: 230, loss: 0.005236648488789797
step: 240, loss: 0.02600875124335289
step: 250, loss: 0.00644551170989871
step: 260, loss: 0.029445167630910873
step: 270, loss: 0.008709128014743328
step: 280, loss: 0.08381425589323044
step: 290, loss: 0.10033876448869705
step: 300, loss: 0.047171831130981445
```

```
step: 310, loss: 0.24894152581691742
step: 320, loss: 0.18639370799064636
step: 330, loss: 0.007829047739505768
step: 340, loss: 0.024330396205186844
step: 350, loss: 0.01764943264424801
step: 360, loss: 0.007674126420170069
step: 370, loss: 0.04724426940083504
step: 380, loss: 0.03333905711770058
step: 390, loss: 0.015115825459361076
step: 400, loss: 0.01922222599387169
step: 410, loss: 0.02120160683989525
step: 420, loss: 0.06335967779159546
step: 430, loss: 0.06756192445755005
step: 440, loss: 0.08309868723154068
step: 450, loss: 0.02542690746486187
step: 460, loss: 0.13960033655166626
step: 470, loss: 0.005960463546216488
step: 480, loss: 0.023528987541794777
step: 490, loss: 0.02865196019411087
step: 500, loss: 0.12445604801177979
step: 510, loss: 0.006914323195815086
step: 520, loss: 0.007543391548097134
step: 530, loss: 0.0409812405705452
/usr/local/lib/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=0.9252767527675276, f1=0.922089825847846,
best_f1=0.922089825847846
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019209954189136624
step: 10, loss: 0.004972134716808796
step: 20, loss: 0.012011205777525902
step: 30, loss: 0.0975329726934433
step: 40, loss: 0.013108761981129646
step: 50, loss: 0.030244771391153336
step: 60, loss: 0.015060166828334332
step: 70, loss: 0.001329674618318677
step: 80, loss: 0.02193155512213707
step: 90, loss: 0.09655134379863739
step: 100, loss: 0.15269030630588531
step: 110, loss: 0.004047837574034929
step: 120, loss: 0.0008279731264337897
```

```
step: 130, loss: 0.24221602082252502
step: 140, loss: 0.031233929097652435
step: 150, loss: 0.001955186016857624
step: 160, loss: 0.004215138498693705
step: 170, loss: 0.02324003539979458
step: 180, loss: 0.012672719545662403
step: 190, loss: 0.006292663514614105
step: 200, loss: 0.00136555265635252
step: 210, loss: 0.06495816260576248
step: 220, loss: 0.004069698508828878
step: 230, loss: 0.18044939637184143
step: 240, loss: 0.008829194121062756
step: 250, loss: 0.006917056627571583
step: 260, loss: 0.10548368841409683
step: 270, loss: 0.0344894602894783
step: 280, loss: 0.005829062312841415
step: 290, loss: 0.06787889450788498
step: 300, loss: 0.00896257534623146
step: 310, loss: 0.06189412623643875
step: 320, loss: 0.23205359280109406
step: 330, loss: 0.03919725492596626
step: 340, loss: 0.08300520479679108
step: 350, loss: 0.2133614420890808
step: 360, loss: 0.025033822283148766
step: 370, loss: 0.05962863191962242
step: 380, loss: 0.06415178626775742
step: 390, loss: 0.07075861841440201
step: 400, loss: 0.008592581376433372
step: 410, loss: 0.020245598629117012
step: 420, loss: 0.024902451783418655
step: 430, loss: 0.09985455125570297
step: 440, loss: 0.1858007162809372
step: 450, loss: 0.006587910931557417
step: 460, loss: 0.010417385026812553
step: 470, loss: 0.0028303361032158136
step: 480, loss: 0.13007113337516785
step: 490, loss: 0.048248905688524246
step: 500, loss: 0.03562309592962265
step: 510, loss: 0.030047880485653877
step: 520, loss: 0.01829318143427372
step: 530, loss: 0.0025093117728829384
/usr/local/lib/python3.7/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=0.9195509822263798, f1=0.9237918215613383,
best_f1=0.922089825847846
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009186256676912308
step: 10, loss: 0.01509370282292366
step: 20, loss: 0.0009173143189400434
step: 30, loss: 0.014731616713106632
step: 40, loss: 0.045543450862169266
step: 50, loss: 0.0172798540443182
step: 60, loss: 0.003289037849754095
step: 70, loss: 0.0012656993931159377
step: 80, loss: 0.001456470345146954
step: 90, loss: 0.019230099394917488
step: 100, loss: 0.0009626815444789827
step: 110, loss: 0.0012171383714303374
step: 120, loss: 0.00406778696924448
step: 130, loss: 0.0012833827640861273
step: 140, loss: 0.002062883460894227
step: 150, loss: 0.001453302800655365
step: 160, loss: 0.012864760123193264
step: 170, loss: 0.01919698901474476
step: 180, loss: 0.005061700474470854
step: 190, loss: 0.0008391771116293967
step: 200, loss: 0.0007962446543388069
step: 210, loss: 0.0038789755199104548
step: 220, loss: 0.003437336767092347
step: 230, loss: 0.0013160865055397153
step: 240, loss: 0.04239313676953316
step: 250, loss: 0.0048691341653466225
step: 260, loss: 0.10167287290096283
step: 270, loss: 0.018082506954669952
step: 280, loss: 0.05782086029648781
step: 290, loss: 0.0004436296003405005
step: 300, loss: 0.21791890263557434
step: 310, loss: 0.00039655837463214993
step: 320, loss: 0.0026532174088060856
step: 330, loss: 0.005069948732852936
step: 340, loss: 0.0006806535529904068
step: 350, loss: 0.001960809575393796
step: 360, loss: 0.002522827358916402
step: 370, loss: 0.0012497123098000884
step: 380, loss: 0.05272922292351723
step: 390, loss: 0.004345565568655729
step: 400, loss: 0.01015467382967472
step: 410, loss: 0.0024608152452856302
step: 420, loss: 0.0031187019776552916
step: 430, loss: 0.07065431773662567
step: 440, loss: 0.008929132483899593
step: 450, loss: 0.006418267264962196
step: 460, loss: 0.003655265085399151
```

```
step: 470, loss: 0.010900053195655346
step: 480, loss: 0.0023287830408662558
step: 490, loss: 0.03619292378425598
step: 500, loss: 0.049223750829696655
step: 510, loss: 0.0010940569918602705
step: 520, loss: 0.005679595284163952
step: 530, loss: 0.05370074138045311
/usr/local/lib/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=0.9269427640763147, f1=0.9203132197144173,
best_f1=0.9203132197144173
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013606674037873745
step: 10, loss: 0.010835414752364159
step: 20, loss: 0.0032526704017072916
step: 30, loss: 0.042496539652347565
step: 40, loss: 0.010953984223306179
step: 50, loss: 0.0076142363250255585
step: 60, loss: 0.007807970512658358
step: 70, loss: 0.00496419332921505
step: 80, loss: 0.0009935119887813926
step: 90, loss: 0.0006764019490219653
step: 100, loss: 0.0010983243118971586
step: 110, loss: 0.012682152912020683
step: 120, loss: 0.05622203439474106
step: 130, loss: 0.012350800447165966
step: 140, loss: 0.006420357618480921
step: 150, loss: 0.07061058282852173
step: 160, loss: 0.0009206852992065251
step: 170, loss: 0.0025006942451000214
step: 180, loss: 0.005768869072198868
step: 190, loss: 0.07265761494636536
step: 200, loss: 0.006296905223280191
step: 210, loss: 0.006537828128784895
step: 220, loss: 0.008273574523627758
step: 230, loss: 0.0005000120145268738
step: 240, loss: 0.00207877391949296
step: 250, loss: 0.0011392912128940225
step: 260, loss: 0.01042575016617775
step: 270, loss: 0.0013558984501287341
step: 280, loss: 0.018601400777697563
```

```
step: 290, loss: 0.018085908144712448
step: 300, loss: 0.001301307580433786
step: 310, loss: 0.005119273904711008
step: 320, loss: 0.04284349083900452
step: 330, loss: 0.007713310420513153
step: 340, loss: 0.003702232614159584
step: 350, loss: 0.0298578143119812
step: 360, loss: 0.0009473293903283775
step: 370, loss: 0.03279055282473564
step: 380, loss: 0.0065380120649933815
step: 390, loss: 0.0013970669824630022
step: 400, loss: 0.04285089671611786
step: 410, loss: 0.025767561048269272
step: 420, loss: 0.08175572007894516
step: 430, loss: 0.0016607866855338216
step: 440, loss: 0.09432480484247208
step: 450, loss: 0.006447803694754839
step: 460, loss: 0.04733172804117203
step: 470, loss: 0.1292523443698883
step: 480, loss: 0.007049970794469118
step: 490, loss: 0.0056831855326890945
step: 500, loss: 0.0030340568628162146
step: 510, loss: 0.00015224669186864048
step: 520, loss: 0.02749895676970482
step: 530, loss: 0.002412583911791444
/usr/local/lib/python3.7/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=0.9235074626865672, f1=0.9207232267037554,
best_f1=0.9203132197144173
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005107151344418526
step: 10, loss: 0.03817227855324745
step: 20, loss: 0.012795616872608662
step: 30, loss: 0.005241445731371641
step: 40, loss: 0.09598927944898605
step: 50, loss: 0.0005046756123192608
step: 60, loss: 0.006294738035649061
step: 70, loss: 0.10478395968675613
step: 80, loss: 0.003594439709559083
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 90, loss: 0.006832101847976446
step: 100, loss: 0.0018935361877083778
step: 110, loss: 0.002968026325106621
step: 120, loss: 0.00034565385431051254
step: 130, loss: 0.09894755482673645
```

```
step: 140, loss: 0.001399313798174262
step: 150, loss: 0.003572588786482811
step: 160, loss: 0.001356308814138174
step: 170, loss: 0.012790828011929989
step: 180, loss: 0.003612534375861287
step: 190, loss: 0.00029808279941789806
step: 200, loss: 0.020203329622745514
step: 210, loss: 0.0011520389234647155
step: 220, loss: 0.0004530186124611646
step: 230, loss: 0.0007932488224469125
step: 240, loss: 0.00704659894108723
step: 250, loss: 0.001372624421492219
step: 260, loss: 0.002181756542995572
step: 270, loss: 0.000591650721617043
step: 280, loss: 0.012392930686473846
step: 290, loss: 0.04438130557537079
step: 300, loss: 0.001060502021573484
step: 310, loss: 0.006048810668289661
step: 320, loss: 0.2618841230869293
step: 330, loss: 0.004885224159806967
step: 340, loss: 0.01578442007303238
step: 350, loss: 0.0008102601859718561
step: 360, loss: 0.0022313783410936594
step: 370, loss: 0.08369754999876022
step: 380, loss: 0.23950275778770447
step: 390, loss: 0.0002196093846578151
step: 400, loss: 0.0006807508761994541
step: 410, loss: 0.0028516713064163923
step: 420, loss: 0.002082433318719268
step: 430, loss: 0.0002555630635470152
step: 440, loss: 0.0002038502716459334
step: 450, loss: 0.0005141685251146555
step: 460, loss: 0.0035659312270581722
step: 470, loss: 0.0032298932783305645
step: 480, loss: 0.000546599505469203
step: 490, loss: 0.0005727086099795997
step: 500, loss: 6.332264456432313e-05
step: 510, loss: 0.0225521232932806
step: 520, loss: 0.010619569569826126
step: 530, loss: 0.0051314192824065685
/usr/local/lib/python3.7/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=0.9216330361332709, f1=0.9172510518934082,
best_f1=0.9203132197144173
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.033958885818719864
step: 10, loss: 0.0037672824691981077
step: 20, loss: 0.03074525110423565
step: 30, loss: 0.013279133476316929
step: 40, loss: 0.0057583097368478775
step: 50, loss: 0.0013712788932025433
step: 60, loss: 0.0006333215278573334
step: 70, loss: 0.020284811662273407
step: 80, loss: 0.0008911509066820145
step: 90, loss: 0.0012969301315024495
step: 100, loss: 0.021696802228689194
step: 110, loss: 0.0007148307049646974
step: 120, loss: 0.004050685092806816
step: 130, loss: 0.0004349604423623532
step: 140, loss: 0.00010805201600305736
step: 150, loss: 0.0015403825091198087
step: 160, loss: 0.0006113757262937725
step: 170, loss: 0.004918523598462343
step: 180, loss: 0.0009703614050522447
step: 190, loss: 0.00136650656349957
step: 200, loss: 0.0024032501969486475
step: 210, loss: 0.002878741594031453
step: 220, loss: 0.003965829499065876
step: 230, loss: 0.0034554244484752417
step: 240, loss: 0.0009813676588237286
step: 250, loss: 0.005998709239065647
step: 260, loss: 0.025215161964297295
step: 270, loss: 0.0005071047344245017
step: 280, loss: 0.000976374140009284
step: 290, loss: 0.008863621391355991
step: 300, loss: 0.000339464342687279
step: 310, loss: 0.02393220365047455
step: 320, loss: 0.00044295049156062305
step: 330, loss: 0.018380625173449516
step: 340, loss: 0.002355682197958231
step: 350, loss: 0.06472907960414886
step: 360, loss: 0.0025877144653350115
step: 370, loss: 0.03140442445874214
step: 380, loss: 0.004159353207796812
step: 390, loss: 0.0009204224916175008
step: 400, loss: 0.025583021342754364
step: 410, loss: 0.08637730032205582
step: 420, loss: 0.007548315450549126
step: 430, loss: 0.00026900041848421097
step: 440, loss: 0.007688338868319988
step: 450, loss: 0.0008813765598461032
step: 460, loss: 0.010619509033858776
step: 470, loss: 0.0012146710650995374
```

```
step: 480, loss: 0.002356824232265353
step: 490, loss: 0.0015109209343791008
step: 500, loss: 0.006121565587818623
step: 510, loss: 0.0004903677036054432
step: 520, loss: 0.0008250465616583824
step: 530, loss: 0.00031677199876867235
/usr/local/lib/python3.7/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=0.9260808926080892, f1=0.9209431345353675,
best_f1=0.9203132197144173
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007408862002193928
step: 10, loss: 0.0006761269178241491
step: 20, loss: 0.00033740472281351686
step: 30, loss: 0.0002869062009267509
step: 40, loss: 0.01926497183740139
step: 50, loss: 0.006623427849262953
step: 60, loss: 0.003040861338376999
step: 70, loss: 0.11203684657812119
step: 80, loss: 0.002127132611349225
step: 90, loss: 0.006225505843758583
step: 100, loss: 0.0003420862485654652
step: 110, loss: 0.00905345194041729
step: 120, loss: 0.0022073432337492704
step: 130, loss: 0.017012959346175194
step: 140, loss: 0.04908117651939392
step: 150, loss: 0.0012133143609389663
step: 160, loss: 0.003404427319765091
step: 170, loss: 0.0009959795279428363
step: 180, loss: 0.00011096005619037896
step: 190, loss: 0.004546568728983402
step: 200, loss: 0.10908973962068558
step: 210, loss: 0.00019184288976248354
step: 220, loss: 0.0004070436698384583
step: 230, loss: 0.012133534997701645
step: 240, loss: 0.00027232145657762885
step: 250, loss: 0.18102316558361053
step: 260, loss: 0.000859725521877408
step: 270, loss: 0.005721749272197485
step: 280, loss: 0.006081827450543642
step: 290, loss: 7.301072037080303e-05
step: 300, loss: 0.00012276771303731948
step: 310, loss: 0.00010992186435032636
step: 320, loss: 6.974356074351817e-05
step: 330, loss: 0.05466211959719658
```

```
step: 340, loss: 0.00041496625635772943
step: 350, loss: 0.000186371777090244
step: 360, loss: 0.05805649980902672
step: 370, loss: 0.0005359448259696364
step: 380, loss: 7.893162546679378e-05
step: 390, loss: 0.0002832758182194084
step: 400, loss: 0.012148820795118809
step: 410, loss: 0.0009492803364992142
step: 420, loss: 0.0024452409707009792
step: 430, loss: 0.005866818595677614
step: 440, loss: 0.0025742880534380674
step: 450, loss: 0.0004510578000918031
step: 460, loss: 0.00014652777463197708
step: 470, loss: 0.00031665401184000075
step: 480, loss: 0.022494519129395485
step: 490, loss: 0.0027011409401893616
step: 500, loss: 0.0033927273470908403
step: 510, loss: 0.08491724729537964
step: 520, loss: 4.6566667151637375e-05
step: 530, loss: 8.172159868991002e-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=0.9245810055865922, f1=0.9193697868396664,
best_f1=0.9203132197144173
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010743503225967288
step: 10, loss: 0.00067195855508108139
step: 20, loss: 0.0005650680977851152
step: 30, loss: 0.004172361455857754
step: 40, loss: 0.00037908810190856457
step: 50, loss: 0.0067272973246872425
step: 60, loss: 0.0024306094273924828
step: 70, loss: 0.00723270233720541
step: 80, loss: 0.003233222523704171
step: 90, loss: 0.0009264893596991897
step: 100, loss: 0.0005035924259573221
step: 110, loss: 0.0009446549811400473
step: 120, loss: 0.002581080188974738
step: 130, loss: 0.00010063420631922781
step: 140, loss: 0.001823881291784346
step: 150, loss: 0.00039372630999423563
step: 160, loss: 6.158564065117389e-05
step: 170, loss: 0.0003544404171407223
step: 180, loss: 0.0030987814534455538
step: 190, loss: 0.00018916615226771683
```

```
step: 200, loss: 0.0001690169592620805
step: 210, loss: 0.0009589603287167847
step: 220, loss: 0.0007303986931219697
step: 230, loss: 0.0002485353616066277
step: 240, loss: 0.0004936606856063008
step: 250, loss: 0.007861967198550701
step: 260, loss: 0.007083934731781483
step: 270, loss: 0.0005458305822685361
step: 280, loss: 0.00023193028755486012
step: 290, loss: 0.0015768844168633223
step: 300, loss: 0.0004964789259247482
step: 310, loss: 0.0020047714933753014
step: 320, loss: 0.0002714372531045228
step: 330, loss: 0.002448488026857376
step: 340, loss: 8.107949543045834e-05
step: 350, loss: 8.295872976304963e-05
step: 360, loss: 4.975170304533094e-05
step: 370, loss: 0.003792399540543556
step: 380, loss: 0.00023309685639105737
step: 390, loss: 0.00010716268297983333
step: 400, loss: 0.00011469150922494009
step: 410, loss: 6.435982504626736e-05
step: 420, loss: 0.0003660245565697551
step: 430, loss: 0.00015479442663490772
step: 440, loss: 7.568877481389791e-05
step: 450, loss: 0.0003952413680963218
step: 460, loss: 0.002002328634262085
step: 470, loss: 0.00015392826753668487
step: 480, loss: 0.0007228507893159986
step: 490, loss: 0.05769513174891472
step: 500, loss: 0.005692363251000643
step: 510, loss: 0.016541069373488426
step: 520, loss: 0.00509223947301507
step: 530, loss: 0.0043909489177167416
/usr/local/lib/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=0.9277389277389277, f1=0.9190189726978251,
best_f1=0.9190189726978251
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013964172103442252
step: 10, loss: 0.002861024346202612
```

```
step: 20, loss: 0.00022160385560709983
step: 30, loss: 0.00021771877072751522
step: 40, loss: 0.0012963837943971157
step: 50, loss: 0.010015456937253475
step: 60, loss: 0.0005034729838371277
step: 70, loss: 0.000363768805982545
step: 80, loss: 0.00018715605256147683
step: 90, loss: 0.003459273837506771
step: 100, loss: 0.002002905821427703
step: 110, loss: 0.0006907240021973848
step: 120, loss: 0.01257980614900589
step: 130, loss: 0.0007786608766764402
step: 140, loss: 0.0002163688332075253
step: 150, loss: 0.00048825526027940214
step: 160, loss: 0.02100525051355362
step: 170, loss: 0.002792270388454199
step: 180, loss: 3.0457264074357226e-05
step: 190, loss: 0.02114463970065117
step: 200, loss: 0.00041850286652334034
step: 210, loss: 0.024950139224529266
step: 220, loss: 0.003375513246282935
step: 230, loss: 0.0009140897891484201
step: 240, loss: 2.9857337722205557e-05
step: 250, loss: 0.0016318231355398893
step: 260, loss: 0.0013426674995571375
step: 270, loss: 0.0022368396166712046
step: 280, loss: 0.0015772937331348658
step: 290, loss: 0.002068268833681941
step: 300, loss: 0.049203552305698395
step: 310, loss: 0.0074387709610164165
step: 320, loss: 0.008122345432639122
step: 330, loss: 0.00047248159535229206
step: 340, loss: 0.00012392942153383046
step: 350, loss: 0.004097200930118561
step: 360, loss: 0.0003422778681859235
step: 370, loss: 5.676318323821761e-05
step: 380, loss: 9.066964412340894e-05
step: 390, loss: 0.006123158149421215
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 400, loss: 5.586988845607266e-05
step: 410, loss: 0.03664693236351013
step: 420, loss: 0.021842053160071373
step: 430, loss: 0.0007198340608738363
step: 440, loss: 0.00025831133825704455
step: 450, loss: 0.00021954037947580218
step: 460, loss: 0.014052625745534897
step: 470, loss: 0.021240482106804848
step: 480, loss: 0.00025466945953667164
```

```
step: 490, loss: 0.0027063353918492794
step: 500, loss: 9.503343608230352e-05
step: 510, loss: 4.373306728666648e-05
step: 520, loss: 5.749621777795255e-05
step: 530, loss: 0.00032386701786890626
/usr/local/lib/python3.7/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=0.9271644525881814, f1=0.9204906860517946,
best_f1=0.9190189726978251
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007841188926249743
step: 10, loss: 0.000893199525307864
step: 20, loss: 2.5607167117414065e-05
step: 30, loss: 6.239274807740003e-05
step: 40, loss: 4.1101055103354156e-05
step: 50, loss: 5.6521912483731285e-05
step: 60, loss: 0.0007556125638075173
step: 70, loss: 0.052960604429244995
step: 80, loss: 0.0007303970633074641
step: 90, loss: 8.514926594216377e-05
step: 100, loss: 0.0013562663225457072
step: 110, loss: 0.0023522761184722185
step: 120, loss: 0.0006879505817778409
step: 130, loss: 0.013003222644329071
step: 140, loss: 0.00012602195783983916
step: 150, loss: 0.0025197011418640614
step: 160, loss: 0.00010779185686260462
step: 170, loss: 0.00034019199665635824
step: 180, loss: 0.00017746358935255557
step: 190, loss: 6.715837662341073e-05
step: 200, loss: 7.272326911333948e-05
step: 210, loss: 0.00046547979582101107
step: 220, loss: 9.287888678954914e-05
step: 230, loss: 0.0032758857123553753
step: 240, loss: 0.0006551974220201373
step: 250, loss: 0.00201993715018034
step: 260, loss: 0.005185683723539114
step: 270, loss: 0.01907942071557045
step: 280, loss: 5.5180269555421546e-05
step: 290, loss: 0.00717940041795373
step: 300, loss: 0.12593376636505127
step: 310, loss: 0.00012991364928893745
step: 320, loss: 0.018316252157092094
step: 330, loss: 8.001859532669187e-05
step: 340, loss: 0.0001608148741070181
```

```
step: 350, loss: 0.014454133808612823
step: 360, loss: 0.00874333456158638
step: 370, loss: 0.0002782406809274107
step: 380, loss: 0.00040431172237731516
step: 390, loss: 3.4264146961504593e-05
step: 400, loss: 9.370953921461478e-05
step: 410, loss: 8.19559209048748e-05
step: 420, loss: 0.026071745902299988
step: 430, loss: 4.531933882390149e-05
step: 440, loss: 0.0004702291334979236
step: 450, loss: 0.050088319927453995
step: 460, loss: 0.0005326354876160622
step: 470, loss: 0.10146624594926834
step: 480, loss: 0.00014288430975284427
step: 490, loss: 0.0006241650553420186
step: 500, loss: 0.000902982777915895
step: 510, loss: 0.000229609664529562
step: 520, loss: 0.011922894045710564
step: 530, loss: 0.00020987700554542243
/usr/local/lib/python3.7/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=0.9268518518518517, f1=0.9249084249084248,
best_f1=0.9190189726978251
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003468220354989171
step: 10, loss: 0.0003290194144938141
step: 20, loss: 0.004689905792474747
step: 30, loss: 0.00014319444017019123
step: 40, loss: 0.0006771263433620334
step: 50, loss: 0.00011992772488156334
step: 60, loss: 0.0005826717824675143
step: 70, loss: 0.00010704750457080081
step: 80, loss: 0.00042537329136393964
step: 90, loss: 0.0006062889588065445
step: 100, loss: 4.9213649617740884e-05
step: 110, loss: 0.000648503948468715
step: 120, loss: 0.00020825353567488492
step: 130, loss: 0.0024936464615166187
step: 140, loss: 0.0032260087318718433
step: 150, loss: 0.0002670141984708607
step: 160, loss: 0.029467910528182983
step: 170, loss: 0.0014926467556506395
step: 180, loss: 0.0006549435202032328
step: 190, loss: 0.000482916715554893
step: 200, loss: 0.00047336870920844376
```

```
step: 210, loss: 0.0006106568616814911
step: 220, loss: 0.0009230768191628158
step: 230, loss: 0.15042507648468018
step: 240, loss: 5.8804205764317885e-05
step: 250, loss: 0.03748045489192009
step: 260, loss: 0.0027876910753548145
step: 270, loss: 0.0005322026554495096
step: 280, loss: 0.04985087364912033
step: 290, loss: 8.392598829232156e-05
step: 300, loss: 0.014679744839668274
step: 310, loss: 0.0013394991401582956
step: 320, loss: 0.0003300299576949328
step: 330, loss: 0.0013688253238797188
step: 340, loss: 0.00035887103877030313
step: 350, loss: 0.001381330075673759
step: 360, loss: 0.00036287822877056897
step: 370, loss: 0.0011870631715282798
step: 380, loss: 0.007113089319318533
step: 390, loss: 0.0001286626938963309
step: 400, loss: 8.799411443760619e-05
step: 410, loss: 4.1810504626482725e-05
step: 420, loss: 0.00036723114317283034
step: 430, loss: 0.0018076056148856878
step: 440, loss: 0.004313584417104721
step: 450, loss: 0.07564035058021545
step: 460, loss: 5.208533548284322e-05
step: 470, loss: 0.012315754778683186
step: 480, loss: 0.00048342716763727367
step: 490, loss: 0.0002030827454291284
step: 500, loss: 0.00010278596892021596
step: 510, loss: 0.00047531374730169773
step: 520, loss: 0.0031497282907366753
step: 530, loss: 0.00015365910076070577
/usr/local/lib/python3.7/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=0.9269195189639223, f1=0.9267399267399268,
best_f1=0.9190189726978251
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.755494541721418e-05
step: 10, loss: 0.18155483901500702
step: 20, loss: 0.00045936982496641576
step: 30, loss: 0.0003297192743048072
step: 40, loss: 0.0007838504388928413
step: 50, loss: 0.0005786079564131796
step: 60, loss: 5.8922203606925905e-05
```

```
step: 70, loss: 0.0001374058920191601
step: 80, loss: 8.521950076101348e-05
step: 90, loss: 0.0002735466696321964
step: 100, loss: 0.0054489802569150925
step: 110, loss: 0.00024109985679388046
step: 120, loss: 0.002004725858569145
step: 130, loss: 0.0001293996028834954
step: 140, loss: 0.006761899683624506
step: 150, loss: 0.031530119478702545
step: 160, loss: 0.0008332353318110108
step: 170, loss: 0.000331413495587185
step: 180, loss: 0.00022546708351001143
step: 190, loss: 0.0002999595017172396
step: 200, loss: 8.389454160351306e-05
step: 210, loss: 0.019729817286133766
step: 220, loss: 0.000580747495405376
step: 230, loss: 8.265860378742218e-05
step: 240, loss: 6.955394928809255e-05
step: 250, loss: 0.002694927854463458
step: 260, loss: 6.336087244562805e-05
step: 270, loss: 0.001706691225990653
step: 280, loss: 0.0001385344221489504
step: 290, loss: 0.00014031252067070454
step: 300, loss: 0.0046536545269191265
step: 310, loss: 0.0074873510748147964
step: 320, loss: 0.008310625329613686
step: 330, loss: 0.00018408357573207468
step: 340, loss: 0.0013167868601158261
step: 350, loss: 0.000584685942158103
step: 360, loss: 0.0038528903387486935
step: 370, loss: 0.00014610144717153162
step: 380, loss: 5.4826352425152436e-05
step: 390, loss: 4.63926226075273e-05
step: 400, loss: 2.310383933945559e-05
step: 410, loss: 0.00034048903034999967
step: 420, loss: 0.00011693626584019512
step: 430, loss: 5.884408165002242e-05
step: 440, loss: 7.688249024795368e-05
step: 450, loss: 5.782841981272213e-05
step: 460, loss: 0.06278377026319504
step: 470, loss: 4.273257218301296e-05
step: 480, loss: 7.555243792012334e-05
step: 490, loss: 0.00015784801507834345
step: 500, loss: 8.982292638393119e-05
step: 510, loss: 0.0003918847651220858
step: 520, loss: 6.602773646591231e-05
step: 530, loss: 0.0001849779364420101
/usr/local/lib/python3.7/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=0.927710843373494, f1=0.9280146721687299,
best_f1=0.9190189726978251
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00020751060219481587
step: 10, loss: 0.0008718629833310843
step: 20, loss: 0.0006882917368784547
step: 30, loss: 6.12790827290155e-05
step: 40, loss: 0.00010222005221294239
step: 50, loss: 0.001485426677390933
step: 60, loss: 3.621573341661133e-05
step: 70, loss: 9.548723755870014e-05
step: 80, loss: 3.372783612576313e-05
step: 90, loss: 2.1416371964733116e-05
step: 100, loss: 0.00023488400620408356
step: 110, loss: 0.12419992685317993
step: 120, loss: 0.00032375851878896356
step: 130, loss: 0.0006084074266254902
step: 140, loss: 4.0728336898609996e-05
step: 150, loss: 0.03959004208445549
step: 160, loss: 0.0006637619808316231
step: 170, loss: 8.248606172855943e-05
step: 180, loss: 0.0007976762135513127
step: 190, loss: 9.025523468153551e-05
step: 200, loss: 0.0007123242248781025
step: 210, loss: 0.0006881073350086808
step: 220, loss: 0.010984978638589382
step: 230, loss: 0.0027084355242550373
step: 240, loss: 0.0003848173073492944
step: 250, loss: 9.30253736441955e-05
step: 260, loss: 0.00015434095985256135
step: 270, loss: 0.005441643297672272
step: 280, loss: 0.00020255280833225697
step: 290, loss: 0.0025365923065692186
step: 300, loss: 0.013277593068778515
step: 310, loss: 0.0019167717546224594
step: 320, loss: 0.0008697222219780087
step: 330, loss: 0.000386200612410903
step: 340, loss: 0.002188002923503518
step: 350, loss: 0.14554893970489502
step: 360, loss: 0.0013266884488984942
step: 370, loss: 0.0011174482060596347
step: 380, loss: 6.382389983627945e-05
step: 390, loss: 0.00030296153272502124
step: 400, loss: 0.000323493528412655
```

```
step: 410, loss: 0.0006994451978243887
step: 420, loss: 8.46257753437385e-05
step: 430, loss: 5.973367296974175e-05
step: 440, loss: 0.002329318318516016
step: 450, loss: 0.005524029023945332
step: 460, loss: 3.528047454892658e-05
step: 470, loss: 0.0001633677602512762
step: 480, loss: 0.0029771693516522646
step: 490, loss: 0.00372058572476244
step: 500, loss: 8.56964397826232e-05
step: 510, loss: 0.0003285503189545125
step: 520, loss: 0.0032425245735794306
step: 530, loss: 0.000268873933237046
/usr/local/lib/python3.7/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=0.927536231884058, f1=0.9274156264447527,
best_f1=0.9190189726978251
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.")
```

```
5742it [00:21, 264.21it/s]
load_f1 = 0.9182915506035284
real_f1 = 0.9168975069252078
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, 268.14it/s]
```