

fastNLP使用FAQ

整理人：许一格， ygxu18@fudan.edu.cn

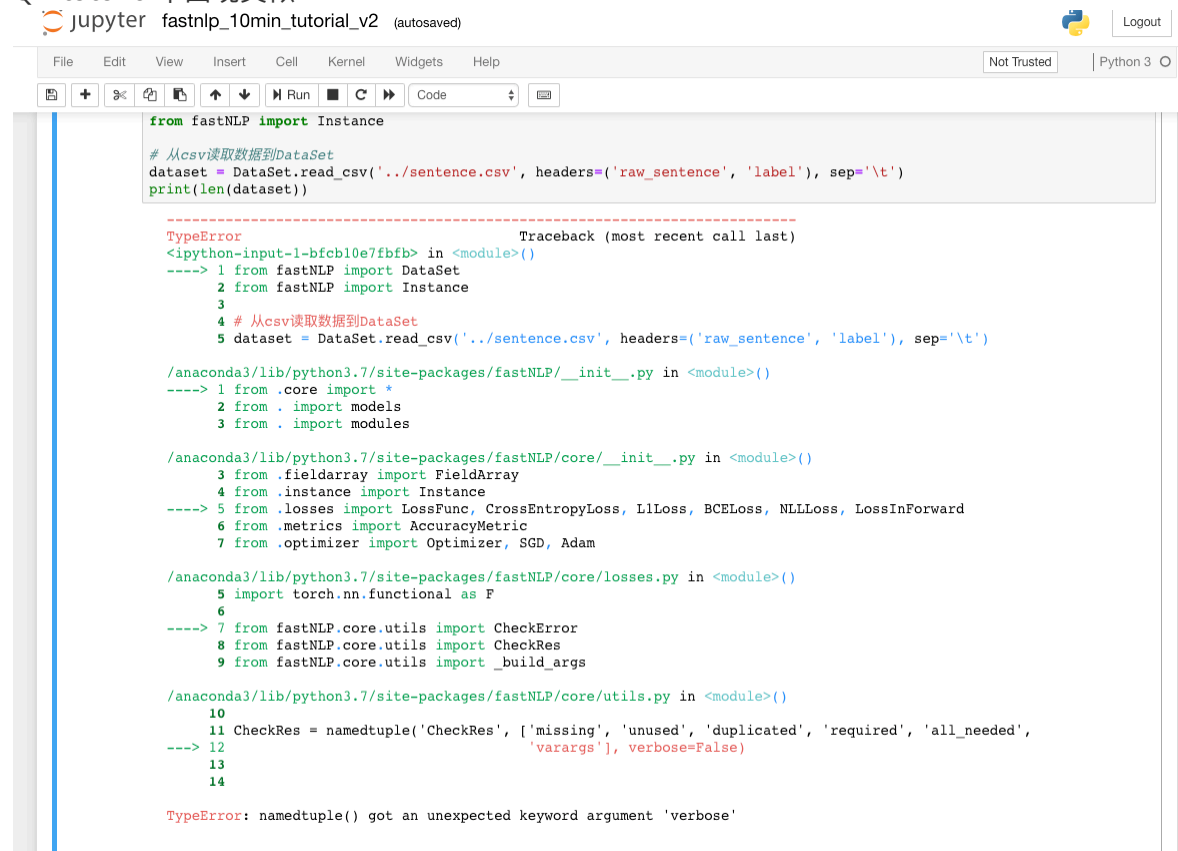
1. Q：怎么安装fastNLP?

A：使用pip安装或者git clone指令

```
pip install fastNLP==0.2.0 -i https://pypi.tuna.tsinghua.edu.cn/simple
```

```
git clone https://github.com/fastnlp/fastNLP.git
```

2. Q：tutorial中出现类似



```
jupyter fastnlp_10min_tutorial_v2 (autosaved) Logout
```

```
File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3
```

```
from fastNLP import Instance

# 从csv读取数据到DataSet
dataset = DataSet.read_csv('../sentence.csv', headers=('raw_sentence', 'label'), sep='\t')
print(len(dataset))

-----
TypeError                                Traceback (most recent call last)
<ipython-input-1-bfcb10e7fbfb> in <module>()
----> 1 from fastNLP import DataSet
      2 from fastNLP import Instance
      3
      4 # 从csv读取数据到DataSet
      5 dataset = DataSet.read_csv('../sentence.csv', headers=('raw_sentence', 'label'), sep='\t')

/anaconda3/lib/python3.7/site-packages/fastNLP/__init__.py in <module>()
----> 1 from .core import *
      2 from . import models
      3 from . import modules

/anaconda3/lib/python3.7/site-packages/fastNLP/core/__init__.py in <module>()
      3 from .fieldarray import FieldArray
      4 from .instance import Instance
----> 5 from .losses import LossFunc, CrossEntropyLoss, L1Loss, BCELoss, NLLLoss, LossInForward
      6 from .metrics import AccuracyMetric
      7 from .optimizer import Optimizer, SGD, Adam

/anaconda3/lib/python3.7/site-packages/fastNLP/core/losses.py in <module>()
      5 import torch.nn.functional as F
      6
----> 7 from fastNLP.core.utils import CheckError
      8 from fastNLP.core.utils import CheckRes
      9 from fastNLP.core.utils import _build_args

/anaconda3/lib/python3.7/site-packages/fastNLP/core/utils.py in <module>()
     10
     11 CheckRes = namedtuple('CheckRes', ['missing', 'unused', 'duplicated', 'required', 'all_needed',
----> 12                                'varargs'], verbose=False)
     13
     14

TypeError: namedtuple() got an unexpected keyword argument 'verbose'
```

的问题时怎么处理?

A： 1) 将python版本从3.7或者3.5更换为3.6

或者

2) 打开fastNLP/core/utils.py第11-12行，将最后的 `verbose=False` 删掉

3. Q：tutorial中出现类似

```

In [4]: from fastNLP import DataSet
        from fastNLP import Instance

        # 从csv读取数据到DataSet
        dataset = DataSet.read_csv('../sentence.csv', headers=('raw_sentence', 'label'), sep='\t')
        print(len(dataset))
        ebook.tqdm in notebook mode. Use tqdm.tqdm instead to force console mode (e.g. in jupyter console)
        " (e.g. in jupyter console)", TqdmExperimentalWarning)

-----
FileNotFoundError                                Traceback (most recent call last)
<ipython-input-4-bfcb10e7fbfb> in <module>()
      3
      4 # 从csv读取数据到DataSet
----> 5 dataset = DataSet.read_csv('../sentence.csv', headers=('raw_sentence', 'label'), sep='\t')
      6 print(len(dataset))

/anaconda3/lib/python3.7/site-packages/fastNLP/core/dataset.py in read_csv(cls, csv_path, headers, sep, dropna)
    330
    331     """
--> 332     with open(csv_path, "r") as f:
    333         start_idx = 0
    334         if headers is None:

FileNotFoundError: [Errno 2] No such file or directory: '../sentence.csv'

```

的问题时怎么处理？

A: tutorial文件夹没有该文件，需要将https://github.com/fastnlp/fastNLP/blob/master/test/data_for_tests/tutorial_sample_dataset.csv下载下来后，更名为sentence.csv，放到指定目录下。

4. Q: tutorial第一句 `from fastNLP import DataSet` 失败了怎么办？

A: 用 `pip list` 命令查看fastNLP包的版本，确保版本号≥0.2.0

5. Q: 按照tutorial的形式调用交叉熵损失函数或者NLL损失函数的时候出错怎么办？

A: 可能是输入损失函数的参数类型有误或者张量的维度不对。需要检查：1) 在 `DataSet().read_csv()` 函数中，默认读入的内容都是字符串，检查label在进入trainer前是否被转换成int；2) 如果传入trainer的是自己写的model，请检测模型返回值是否与fastNLP兼容，以及检测模型返回的张量维度是否正确。

6. Q: 传入fastNLP的模型需要返回什么内容？

A: 返回一个dict。如果在实例化trainer的时候传入了损失函数（与tutorial一样），那么模型返回的dict必须包含一个key为"pred"、value为预测结果的item；如果在实例化trainer的时候没有传入损失函数，那么模型返回的dict必须包含一个key为"loss"、value为loss值的item（即在模型在forward函数当中计算loss值）。

7. Q: 在使用fastNLP的时候怎么定义自己的模型？

A: fastNLP目前兼容pytorch的神经网络模型。模型forward函数的参数名与dataset的field名字一一对应，模型返回值为一个dict，供trainer后续调用，这个dict的组织方法见FAQ第6问。

注: fastNLP在inference的时候会默认调用模型的predict函数，如果模型找不到predict函数则会直接调用forward函数。

8. Q: fastNLP怎么加载预训练的word embedding？

A: 首先import EmbedLoader模块：

```
from fastNLP.io.embed_loader import EmbedLoader
```

然后调用

```
EmbedLoader().load_embedding(emb_dim, emb_file, emb_type, vocab)
```

其中每一个参数的含义如下：

emb_dim: int, 预训练embedding的维度

emb_file: str, 预训练embedding的文件路径

emb_type: str, 预训练embedding的形式, 目前暂时只支持glove (即文件的每一行有emb_dim+1列, 第一列为单词, 后面emb_dim列为这个单词的embedding向量; 文件**没有**文件头)

9. Q: fastNLP目前提供了什么神经网络模块?

A: 提供了Embedding、Conv、Linear、LSTM、Transformer、attention、self_attention、pooling、MLP、CRF等, 详情见fastNLP.modules

更新于2018.12.22