```python
import tensorflow as tf
import ktrain
import pandas as pd
```

```python
#!pip install ktrain
```

```
Collecting keras-pos-embd>=0.11.0
  Downloading https://files.pythonhosted.org/packages/09/70/b63ed8fc660da2bb6ae29b
Collecting keras-multi-head>=0.27.0
  Downloading https://files.pythonhosted.org/packages/e6/32/45adf2549450aca7867dec
Collecting keras-layer-normalization>=0.14.0
  Downloading https://files.pythonhosted.org/packages/a4/0e/d1078df0494bac9ce1a679
Collecting keras-position-wise-feed-forward>=0.6.0
  Downloading https://files.pythonhosted.org/packages/e3/59/f0faa1037c033059e7e9e7
Collecting keras-embed-sim>=0.8.0
  Downloading https://files.pythonhosted.org/packages/57/ef/61a1e39082c9e1834a2d09
Collecting keras-self-attention==0.46.0
  Downloading https://files.pythonhosted.org/packages/15/6b/c804924a056955fa1f3ff7
Building wheels for collected packages: ktrain, langdetect, syntok, seqeval, keras
  Building wheel for ktrain (setup.py) ... done
  Created wheel for ktrain: filename=ktrain-0.25.3-cp36-none-any.whl size=25276312
  Stored in directory: /root/.cache/pip/wheels/86/cd/9c/ad53b98a3eccf4ddeddccfce99
  Building wheel for langdetect (setup.py) ... done
  Created wheel for langdetect: filename=langdetect-1.0.8-cp36-none-any.whl size=9
  Stored in directory: /root/.cache/pip/wheels/8d/b3/aa/6d99de9f3841d7d3d40a60ea06
  Building wheel for syntok (setup.py) ... done
  Created wheel for syntok: filename=syntok-1.3.1-cp36-none-any.whl size=20918 sha
  Stored in directory: /root/.cache/pip/wheels/51/c6/a4/be1920586c49469846bcd28882
  Building wheel for seqeval (setup.py) ... done
  Created wheel for seqeval: filename=seqeval-0.0.19-cp36-none-any.whl size=9919 s
  Stored in directory: /root/.cache/pip/wheels/8d/1f/bf/1198beceed805a2099060975f6
  Building wheel for keras-bert (setup.py) ... done
  Created wheel for keras-bert: filename=keras_bert-0.86.0-cp36-none-any.whl size=
  Stored in directory: /root/.cache/pip/wheels/66/f0/b1/748128b58562fc9e31b907bb5e
  Building wheel for sacremoses (setup.py) ... done
  Created wheel for sacremoses: filename=sacremoses-0.0.43-cp36-none-any.whl size=
  Stored in directory: /root/.cache/pip/wheels/29/3c/fd/7ce5c3f0666dab31a50123635e
  Building wheel for keras-transformer (setup.py) ... done
  Created wheel for keras-transformer: filename=keras_transformer-0.38.0-cp36-none
  Stored in directory: /root/.cache/pip/wheels/e5/fb/3a/37b2b9326c799aa010ae46a04d
  Building wheel for keras-pos-embd (setup.py) ... done

  Created wheel for keras-pos-embd: filename=keras_pos_embd-0.11.0-cp36-none-any.w
  Stored in directory: /root/.cache/pip/wheels/5b/a1/a0/ce6b1d49ba1a9a76f592e70cf2
  Building wheel for keras-multi-head (setup.py) ... done
  Created wheel for keras-multi-head: filename=keras_multi_head-0.27.0-cp36-none-a
  Stored in directory: /root/.cache/pip/wheels/b5/b4/49/0a0c27dcb93c13af02fea254ff
  Building wheel for keras-layer-normalization (setup.py) ... done
  Created wheel for keras-layer-normalization: filename=keras_layer_normalization-
  Stored in directory: /root/.cache/pip/wheels/54/80/22/a638a7d406fd155e507aa33d70
  Building wheel for keras-position-wise-feed-forward (setup.py) ... done
  Created wheel for keras-position-wise-feed-forward: filename=keras_position_wise
  Stored in directory: /root/.cache/pip/wheels/39/e2/e2/3514fef126a00574b13bc0b9e2
  Building wheel for keras-embed-sim (setup.py) ... done
  Created wheel for keras-embed-sim: filename=keras_embed_sim-0.8.0-cp36-none-any.
  Stored in directory: /root/.cache/pip/wheels/49/45/8b/c111f6cc8bec253e984677de73
  Building wheel for keras-self-attention (setup.py) ... done
  Created wheel for keras-self-attention: filename=keras_self_attention-0.46.0-cp3
  Stored in directory: /root/.cache/pip/wheels/d2/2e/80/fec4c05eb23c8e13b790e26d20
Successfully built ktrain langdetect syntok seqeval keras-bert sacremoses keras-tr
```

```
ERROR: transformers 3.5.1 has requirement sentencepiece==0.1.91, but you'll have s
Installing collected packages: threadpoolctl, scikit-learn, langdetect, cchardet,
  Found existing installation: scikit-learn 0.22.2.post1
    Uninstalling scikit-learn-0.22.2.post1:
      Successfully uninstalled scikit-learn-0.22.2.post1
Successfully installed cchardet-2.1.7 keras-bert-0.86.0 keras-embed-sim-0.8.0 kera
```

```python
df=pd.read_csv(r'spam.csv',engine='python')
```

```python
df.head()
```

	v1	v2	Unnamed: 2	Unnamed: 3	Unnamed: 4
0	ham	Go until jurong point, crazy.. Available only ...	NaN	NaN	NaN
1	ham	Ok lar... Joking wif u oni...	NaN	NaN	NaN
2	spam	Free entry in 2 a wkly comp to win FA Cup fina...	NaN	NaN	NaN
3	ham	U dun say so early hor... U c already then say...	NaN	NaN	NaN

```python
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5572 entries, 0 to 5571
Data columns (total 5 columns):
 #   Column      Non-Null Count  Dtype
---  ------      --------------  -----
 0   v1          5572 non-null   object
 1   v2          5572 non-null   object
 2   Unnamed: 2  50 non-null     object
 3   Unnamed: 3  12 non-null     object
 4   Unnamed: 4  6 non-null      object
dtypes: object(5)
memory usage: 217.8+ KB
```

```python
df.drop(['Unnamed: 2','Unnamed: 3','Unnamed: 4'],axis=1,inplace=True)
```

```python
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5572 entries, 0 to 5571
Data columns (total 2 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   v1      5572 non-null   object
 1   v2      5572 non-null   object
dtypes: object(2)
memory usage: 87.2+ KB
```

```
df[['v1']]
```

	v1
0	ham
1	ham
2	spam
3	ham
4	ham
...	...
5567	spam
5568	ham
5569	ham
5570	ham
5571	ham

5572 rows × 1 columns

```
from ktrain import text as txt
```

```
# load data
(X_train, y_train), (X_test, y_test), preproc = txt.texts_from_df(train_df=df,
                                                  text_column = 'v2',
                                                  label_columns = 'v1',
                                                  maxlen = 500,
                                                  preprocess_mode = 'bert
```

```
downloading pretrained BERT model (uncased_L-12_H-768_A-12.zip)...
[████████████████████████████████████████████████]
extracting pretrained BERT model...
done.

cleanup downloaded zip...
done.

preprocessing train...
language: en
done.
Is Multi-Label? False
preprocessing test...
language: en
done.
```

```
model = txt.text_classifier(name = 'bert',
                            train_data = (X_train, y_train),
                            preproc = preproc)
```

```
Is Multi-Label? False
```

```
    maxlen is 500
    done.
```

```python
learner = ktrain.get_learner(model=model, train_data=(X_train, y_train),
                  val_data = (X_test, y_test),
                  batch_size = 6)
```

```python
learner.fit_onecycle(lr = 2e-5, epochs = 1)
```

```
    begin training using onecycle policy with max lr of 2e-05...
    836/836 [==============================] - 868s 1s/step - loss: 0.1550 - accuracy: 0
    <tensorflow.python.keras.callbacks.History at 0x7f3496e21518>
```

```python
predictor = ktrain.get_predictor(learner.model, preproc)
```

```python
data='You have won a lottery of 5000'
```

```python
predictor.predict(data)
```

```
    'spam'
```