

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import tensorflow as tf
```

```
In [2]: text_train = open('data_train.txt', 'r').readlines()
type(text_train)
```

Out[2]: list

```
In [3]: len(text_train)
```

Out[3]: 2000

```
In [4]: class_train = open('labels_train_original.txt', 'r').readlines()
type(class_train)
```

Out[4]: list

```
In [5]: len(text_train)
```

Out[5]: 2000

Create a DataFrame for training

```
In [6]: train_data = {
    'Sentence': text_train,
    'Class': class_train
}
```

```
In [7]: df_train = pd.DataFrame(train_data)
```

```
In [8]: df_train.head(n=6)
```

Out[8]:

	Sentence	Class
0	the sign in front of the steepled church read ...	News\n
1	lindsey larsen a soprano and samuel ramey the ...	Classifieds\n
2	to the editor sylvia ann hewlett 's book creat...	Opinion\n
3	illinois tool works inc glenview ill a maker o...	News\n
4	to the editor robert schaeffer op ed feb 19 ex...	Opinion\n
5	to the editor being texan i took great interes...	Opinion\n

Create a DataFrame for testing

```
In [9]: text_test = open('data_valid.txt', 'r').readlines()
class_test = open('labels_valid_original.txt', 'r').readlines()
```

```
In [10]: test_data = {
          'Sentence': text_test,
          'Class': class_test
        }
df_test = pd.DataFrame(test_data)
```

Analysis on Train Data

```
In [11]: df_train['Class'].value_counts(normalize=True)
```

```
Out[11]: Classifieds\n    0.2615
News\n    0.2520
Features\n    0.2440
Opinion\n    0.2425
Name: Class, dtype: float64
```

- We do not have class imbalance. So, we can safely use accuracy as a metric for scoring our models.

Using Fast AI (A Neural Network Library)

```
In [12]: !pip install -Uqq fastbook
import fastbook
fastbook.setup_book()
```

```
C:\Users\itzal\anaconda3\lib\site-packages\fastbook\__init__.py:19: UserWarning: Missing `graphviz` - please run `conda install fastbook`
  except ModuleNotFoundError: warn("Missing `graphviz` - please run `conda install fastbook`")
```

```
In [13]: pip install fastbook
```

```
1.2)
Requirement already satisfied: backcall in c:\users\itzal\anaconda3\lib\site-packages (from ipython<8.0,>=7.23.1->ipykernel->nbdev>=0.2.38->fastbook) (0.2.0)
Requirement already satisfied: decorator in c:\users\itzal\anaconda3\lib\site-packages (from ipython<8.0,>=7.23.1->ipykernel->nbdev>=0.2.38->fastbook) (5.1.0)
Requirement already satisfied: prompt-toolkit!=3.0.0,!<3.0.1,<3.1.0,>=2.0.0 in c:\users\itzal\anaconda3\lib\site-packages (from ipython<8.0,>=7.23.1->ipykernel->nbdev>=0.2.38->fastbook) (3.0.20)
Requirement already satisfied: jedi>=0.16 in c:\users\itzal\anaconda3\lib\site-packages (from ipython<8.0,>=7.23.1->ipykernel->nbdev>=0.2.38->fastbook) (0.17.1)
Requirement already satisfied: pickleshare in c:\users\itzal\anaconda3\lib\site-packages (from ipython<8.0,>=7.23.1->ipykernel->nbdev>=0.2.38->fastbook) (0.7.5)

Requirement already satisfied: parso<0.8.0,>=0.7.0 in c:\users\itzal\anaconda3\lib\site-packages (from jedi>=0.16->ipython<8.0,>=7.23.1->ipykernel->nbdev>=0.2.38->fastbook) (0.7.0)
```

In []:

In []: `from fastai.text.all import *`

```
dls = TextDataLoaders.from_df(train_df, text_col='review', label_col='sentiment')  
dls.show_batch(max_n=3)
```