

```
!pip install autoviml
```

```
Collecting autoviml
```

```
  Downloading https://files.pythonhosted.org/packages/ab/5a/b27f0d9a7341ea18f1d8a861c8364acc98897afd1d837ea46e5e6839  
  |████████████████████████████████████████| 122kB 8.1MB/s
```

```
Collecting scikit-learn>=0.23.1
```

```
  Downloading https://files.pythonhosted.org/packages/e2/4c/6111b9a325f29527d7f262e2ee8c730d354b47a728d955e186dacad5  
  |████████████████████████████████████████| 22.2MB 49.6MB/s
```

```
Requirement already satisfied: jupyter in /usr/local/lib/python3.6/dist-packages (from autoviml) (1.0.0)
```

```
Requirement already satisfied: beautifulsoup4 in /usr/local/lib/python3.6/dist-packages (from autoviml) (4.6.3)
```

```
Requirement already satisfied: nltk in /usr/local/lib/python3.6/dist-packages (from autoviml) (3.2.5)
```

```
Collecting emoji
```

```
  Downloading https://files.pythonhosted.org/packages/79/0d/c06a34f306a219fd32b9d0297d999e3551d0ecce8190dce921602db4  
  |████████████████████████████████████████| 133kB 60.2MB/s
```

```
Collecting xgboost>=1.1.1
```

```
  Downloading https://files.pythonhosted.org/packages/2e/57/bf5026701c384decd2b995eb39d86587a103ba4eb26f8a9b1811db08  
  |████████████████████████████████████████| 157.5MB 82kB/s
```

```
Requirement already satisfied: ipython in /usr/local/lib/python3.6/dist-packages (from autoviml) (5.5.0)
```

```
Collecting vaderSentiment
```

```
  Downloading https://files.pythonhosted.org/packages/76/fc/310e16254683c1ed35eeb97386986d6c00bc29df17ce280aed64d555  
  |████████████████████████████████████████| 133kB 58.2MB/s
```

```
Collecting imbalanced-learn>=0.7
```

```
  Downloading https://files.pythonhosted.org/packages/c8/81/8db4d87b03b998fda7c6f835d807c9ae4e3b141f978597b8d7f31600  
  |████████████████████████████████████████| 174kB 60.4MB/s
```

```
Requirement already satisfied: regex in /usr/local/lib/python3.6/dist-packages (from autoviml) (2019.12.20)
```

```
Requirement already satisfied: matplotlib in /usr/local/lib/python3.6/dist-packages (from autoviml) (3.2.2)
```

```
Collecting shap>=0.36.0
```

```
  Downloading https://files.pythonhosted.org/packages/44/20/54381999efe3000f70a7f68af79ba857cfa3f82278ab0e02e6ba1c06  
  |████████████████████████████████████████| 358kB 52.5MB/s
```

```
Collecting reg-resampler
```

```
  Downloading https://files.pythonhosted.org/packages/7b/fe/911c5f0ee409b18ffeff45cda6e7e7f2196fe604feaf6c4f3483ede4
```

```
Collecting catboost
```

```
  Downloading https://files.pythonhosted.org/packages/20/37/bc4e0ddc30c07a96482abf1de7ed1ca54e59bba2026a33bca6d2ef28  
  |████████████████████████████████████████| 65.8MB 42kB/s
```

```
Requirement already satisfied: textblob in /usr/local/lib/python3.6/dist-packages (from autoviml) (0.15.3)
```

```
Requirement already satisfied: seaborn in /usr/local/lib/python3.6/dist-packages (from autoviml) (0.11.1)
```

```
Requirement already satisfied: pandas in /usr/local/lib/python3.6/dist-packages (from autoviml) (1.1.5)
```

```
Collecting threadpoolctl>=2.0.0
```

```
  Downloading https://files.pythonhosted.org/packages/f7/12/ec3f2e203afa394a149911729357aa48affc59c20e2c1c8297a60f33
```

```
Requirement already satisfied: scipy>=0.19.1 in /usr/local/lib/python3.6/dist-packages (from scikit-learn>=0.23.1->au
```

```
Requirement already satisfied: joblib>=0.11 in /usr/local/lib/python3.6/dist-packages (from scikit-learn>=0.23.1->au
```

```
Requirement already satisfied: numpy>=1.13.3 in /usr/local/lib/python3.6/dist-packages (from scikit-learn>=0.23.1->a
Requirement already satisfied: notebook in /usr/local/lib/python3.6/dist-packages (from jupyter->autoviml) (5.3.1)
Requirement already satisfied: nbconvert in /usr/local/lib/python3.6/dist-packages (from jupyter->autoviml) (5.6.1)
Requirement already satisfied: jupyter-console in /usr/local/lib/python3.6/dist-packages (from jupyter->autoviml) (5
Requirement already satisfied: ipykernel in /usr/local/lib/python3.6/dist-packages (from jupyter->autoviml) (4.10.1)
Requirement already satisfied: ipywidgets in /usr/local/lib/python3.6/dist-packages (from jupyter->autoviml) (7.6.3)
Requirement already satisfied: qtconsole in /usr/local/lib/python3.6/dist-packages (from jupyter->autoviml) (5.0.1)
Requirement already satisfied: six in /usr/local/lib/python3.6/dist-packages (from nltk->autoviml) (1.15.0)
Requirement already satisfied: pexpect; sys_platform != "win32" in /usr/local/lib/python3.6/dist-packages (from ipyt
Requirement already satisfied: pygments in /usr/local/lib/python3.6/dist-packages (from ipython->autoviml) (2.6.1)
Requirement already satisfied: prompt-toolkit<2.0.0,>=1.0.4 in /usr/local/lib/python3.6/dist-packages (from ipython-
Requirement already satisfied: simplegeneric>0.8 in /usr/local/lib/python3.6/dist-packages (from ipython->autoviml)
Requirement already satisfied: traitlets>=4.2 in /usr/local/lib/python3.6/dist-packages (from ipython->autoviml) (4.
Requirement already satisfied: decorator in /usr/local/lib/python3.6/dist-packages (from ipython->autoviml) (4.4.2)
Requirement already satisfied: pickleshare in /usr/local/lib/python3.6/dist-packages (from ipython->autoviml) (0.7.5
Requirement already satisfied: setuptools>=18.5 in /usr/local/lib/python3.6/dist-packages (from ipython->autoviml) (
Requirement already satisfied: requests in /usr/local/lib/python3.6/dist-packages (from vaderSentiment->autoviml) (2
Requirement already satisfied: python-dateutil>=2.1 in /usr/local/lib/python3.6/dist-packages (from matplotlib->auto
Requirement already satisfied: ...
```

```
import tensorflow_datasets as tfds
import numpy as np
import pandas as pd
```

```
df=pd.read_csv('/content/drive/MyDrive/dataset/imdb.csv')
df
```

	text	label
0	I always wrote this series off as being a comp...	0
1	1st watched 12/7/2002 - 3 out of 10(Dir-Steve ...	0
2	This movie was so poorly written and directed ...	0
3	The most interesting thing about Miryang (Secr...	1

```
from sklearn.model_selection import train_test_split
train,test=train_test_split(df,test_size=0.25)
```

ADDE This is the kind of picture John Cassiter would 4

```
from autoviml.Auto_NLP import Auto_NLP
```

[illegible]

```

[nltk_data] | Downloading package treebank to /root/nltk_data...
[nltk_data] | Unzipping corpora/treebank.zip.
[nltk_data] | Downloading package twitter_samples to
[nltk_data] | /root/nltk_data...
[nltk_data] | Unzipping corpora/twitter_samples.zip.
[nltk_data] | Downloading package omw to /root/nltk_data...
[nltk_data] | Unzipping corpora/omw.zip.
[nltk_data] | Downloading package wordnet to /root/nltk_data...
[nltk_data] | Unzipping corpora/wordnet.zip.
[nltk_data] | Downloading package wordnet_ic to /root/nltk_data...
[nltk_data] | Unzipping corpora/wordnet_ic.zip.
[nltk_data] | Downloading package words to /root/nltk_data...
[nltk_data] | Unzipping corpora/words.zip.
[nltk_data] | Downloading package maxent_ne_chunker to
[nltk_data] | /root/nltk_data...
[nltk_data] | Unzipping chunkers/maxent_ne_chunker.zip.
[nltk_data] | Downloading package punkt to /root/nltk_data...
[nltk_data] | Unzipping tokenizers/punkt.zip.
[nltk_data] | Downloading package snowball_data to
[nltk_data] | /root/nltk_data...
[nltk_data] | Downloading package averaged_perceptron_tagger to
[nltk_data] | /root/nltk_data...
[nltk_data] | Unzipping taggers/averaged_perceptron_tagger.zip.
[nltk_data] |

```

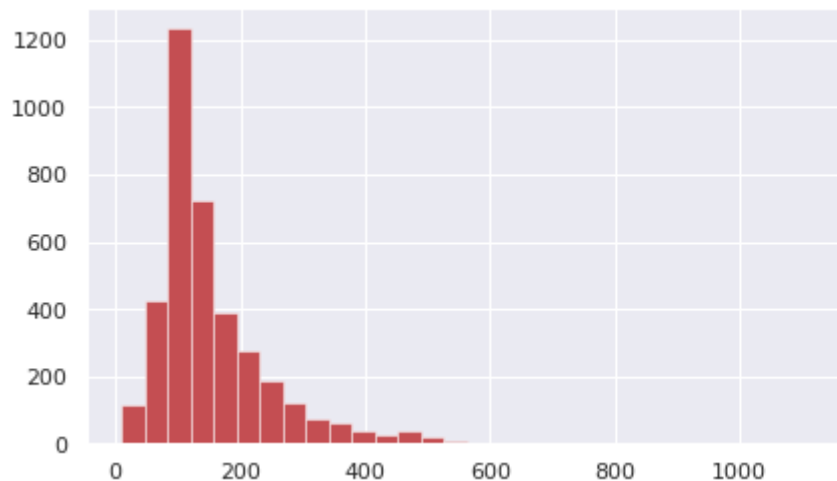
[nltk_data] Done downloading collection popular

Auto NLP processing on NLP Column: text

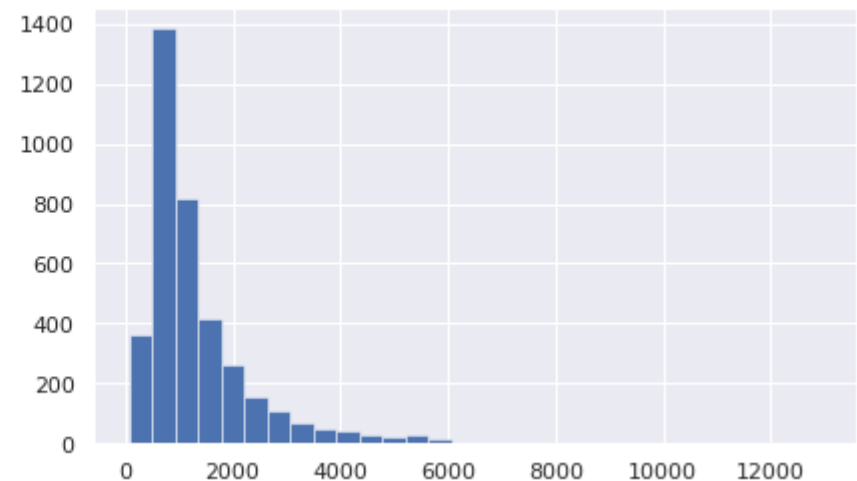
Shape of Train Data: 3750 rows

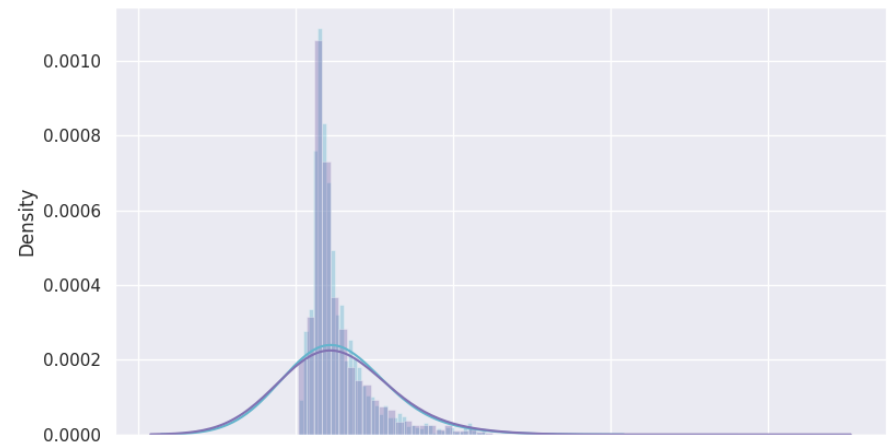
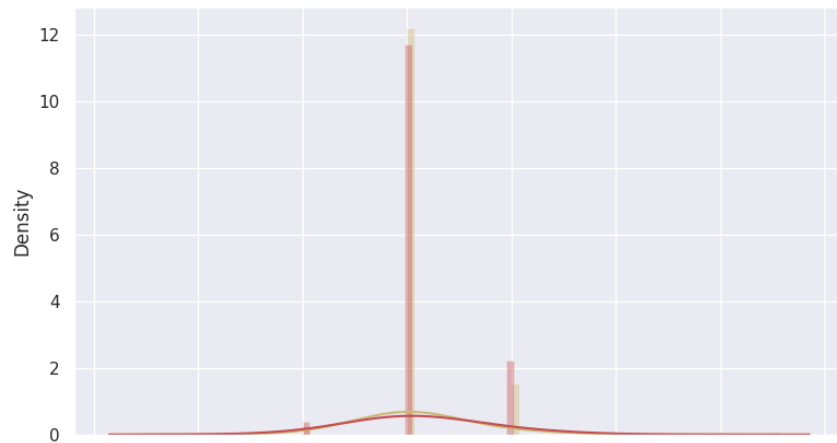
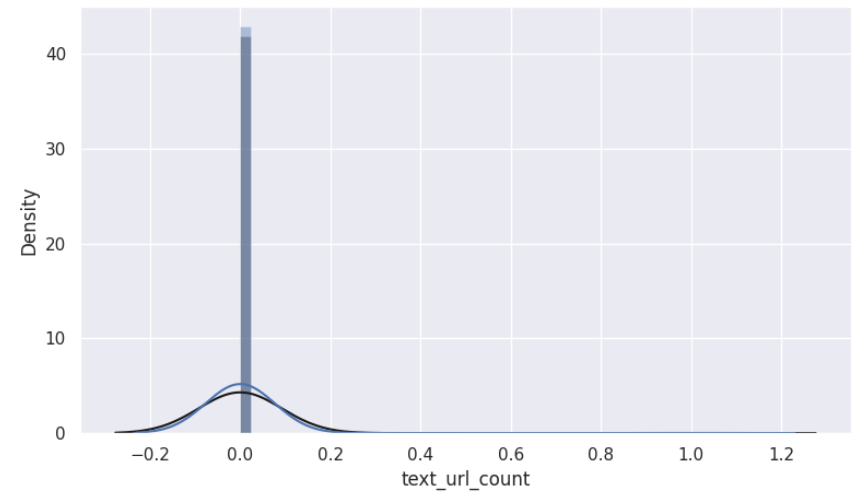
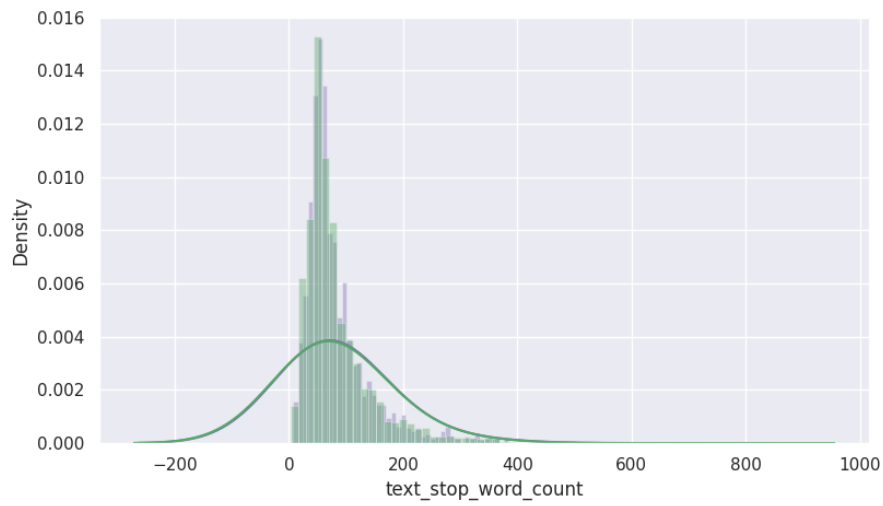
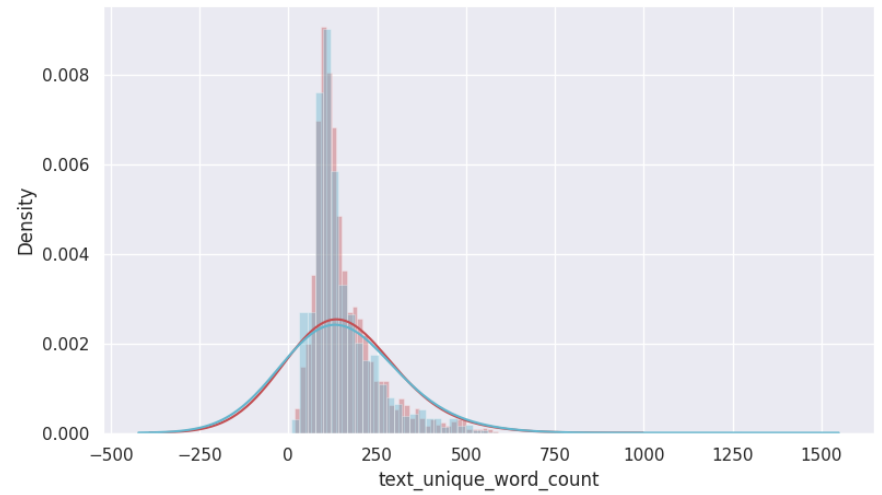
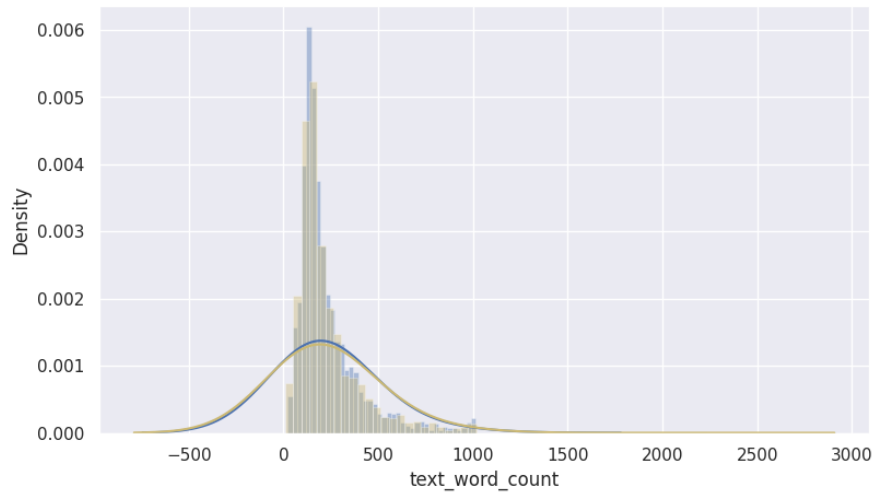
Shape of Test Data: 1250 rows

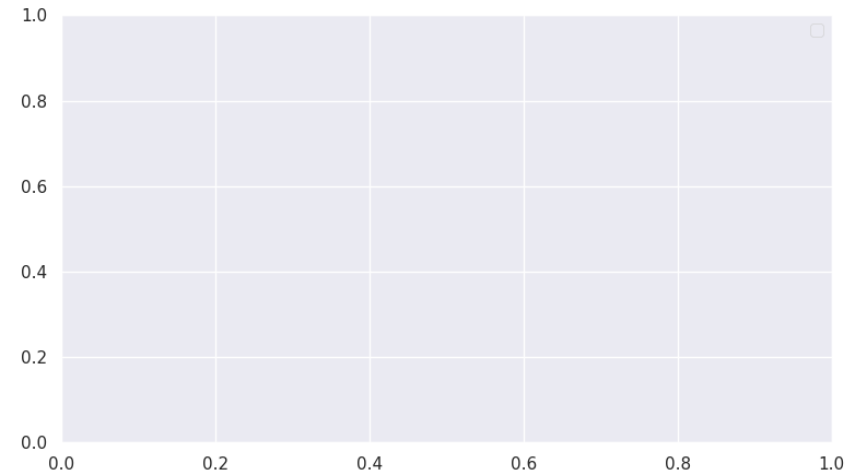
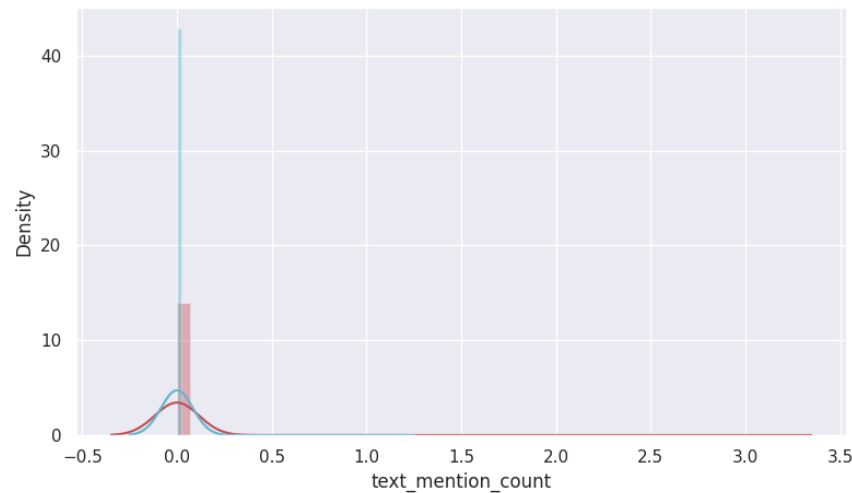
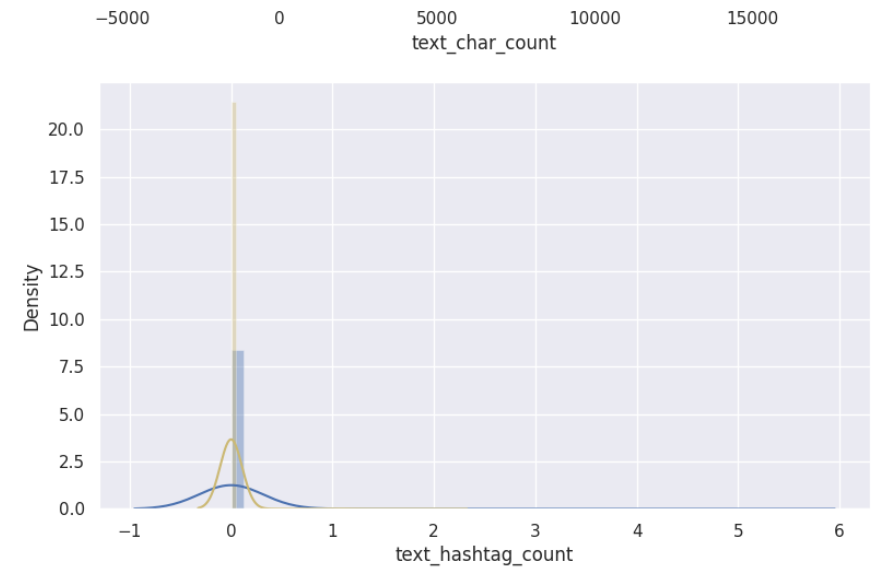
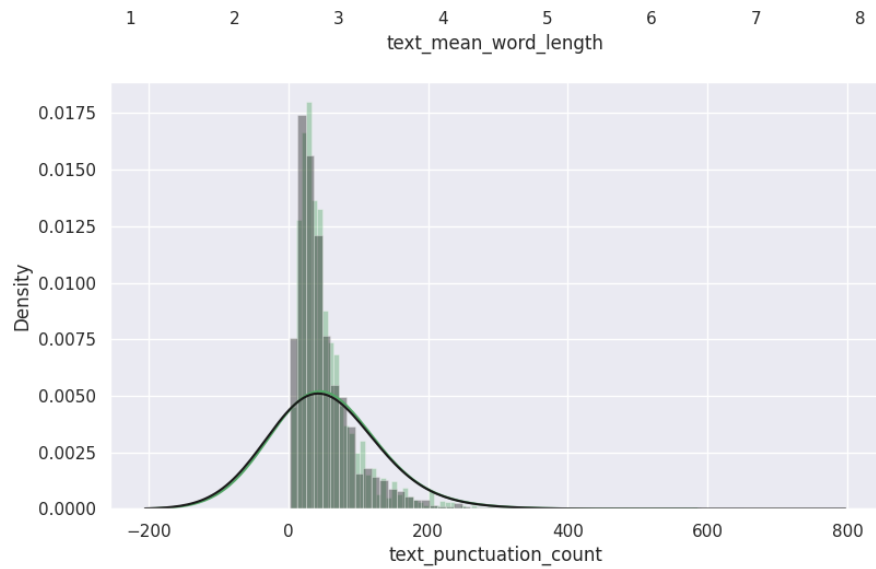
Word Count



Character Count







Added 9 summary columns for counts of words and characters in each row
 Cleaning text in Train data for text column
 Faster text processing using clean_text function, since top_num_features is below 300
 Time Taken for Expanding emojis and abbreviations in data = 23 seconds
 Time Taken for Processing text in data = 43 seconds
 Train data Text cleaning completed. Time taken = 72 seconds

For target = label

For class = 0
 Top 200 n-grams
 : ['had', 'see', 'if', 'good', 'there', 'they', 'watch', 'would', 'time', 'what', 'who', 'even', 'no', 'go', 'think', '']