This task was to predict wine’s variety which applies text classification methods for predicting vaiety from its tasting note, or description.

Dataset Given :

- Train.csv (Renamed Winetrain.csv)

-Test.csv (Renamed as Winetest.csv)

Models Used:

- Linear SVM

- Simple pooling model based on [FastText]

- LSTM

- Bidirectional LSTM

Results:

Test accuracies are below for the 28-class wine variety classification task. As a baseline, a majority class classifier would achieve baseline is 0.13 accuracy (Chardonnay).

| Model | Test Accuracy |

| ------------- | ------------- |

| SVM | 0.75 |

| Simple Pooling | 0.74 |

| LSTM | 0.69 |

| Bi-LSTM | 0.70 |

Word2Vec:

A related task, word embeddings are fit on the wine tasting notes using the Word2Vec algorithm. These reveal interesting clusters and relationships of wine-related words that can be explored.

The saved Bokeh file is given below:

