Sentiment Analysis

170010036, 170010047, 160050012

Precision, Accuracy, Recall, F1 score

```
Classification Report :
              precision    recall  f1-score   support

           1       0.10      0.01      0.02      1271
           2       0.00      0.00      0.00       630
           3       0.07      0.07      0.07       911
           4       0.14      0.46      0.22      1404
           5       0.59      0.45      0.51      5784

    accuracy                           0.33     10000
   macro avg       0.18      0.20      0.16     10000
weighted avg       0.38      0.33      0.34     10000
```

Model Architecture

```
Model: "sequential_5"

_____
Layer (type)                    Output Shape              Param #
=================================================================
dense_5 (Dense)                 (None, 5)                 305
=================================================================
Total params: 305
Trainable params: 305
Non-trainable params: 0
```
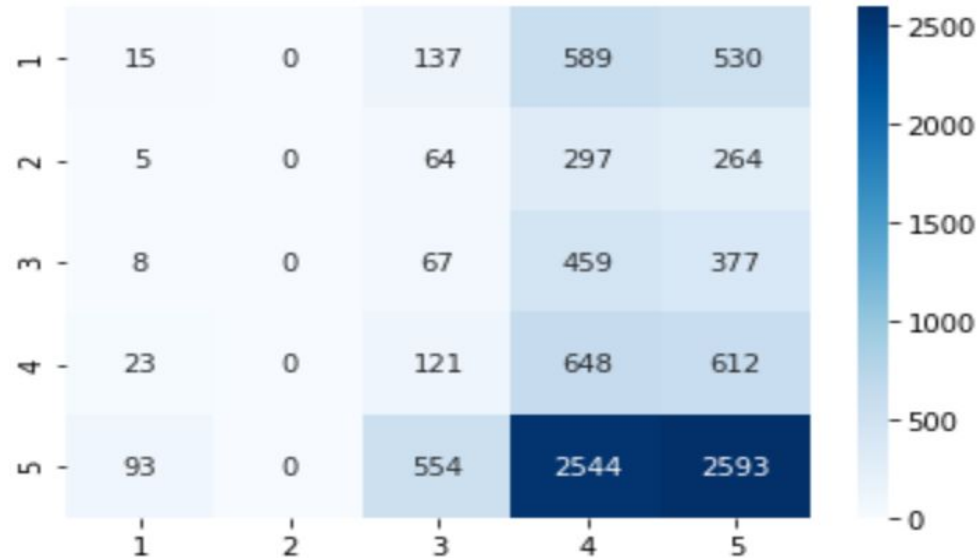
Due to such low number of parameters and simple architecture, the model is giving a very low accuracy as we shown in previous slide.

Heat Map of Confusion Matrix

We can see that our model is predicting 4, 5 most of times. This is because around 65% of the training data consists of class 5. Also, our model architecture is very simple.

Some Examples:

Input: Cards are not as big as pictured.

Output: 3

Input: Do not buy these! They break very fast I spun then for 15 minutes and the end flew off don't waste your money. They are made from cheap plastic and have cracks in them. Buy the poi balls they work a lot better if you only have limited funds.

Output: 1

Input: To keep together, had to use crazy glue

Output: 5

But results are not consistent due to bad distribution of data. For example, if we train the data again, with a very high probability, we will get different results for the same test sentences.. This is due to class imbalance of the data as well as the model architecture.