Used tf-idf to remove stop words and other highly frequent words (for sentences of length greater than 4)

Removed sentences that had no words, as this would cause high bias for that sentiment

Converted words into embeddings

Applied text to sequence

Performed LSTM with 128 units

The results were highly biased towards class that covered maximum area in the distribution

Tried different techniques to balance the dataset

Applied SMOTE

Shuffled dataset

Even after applying various balancing techniques

The results were highly skewed, and the model only predicted single class. Tried options:

Increasing/decreasing learning rate

Changing optimizer

Changing batch-size

By increasing the batch-size significantly the learning improved and started converging towards minima. After certain point, it stopped learning