

Covid-19 Fact Checking

Hambrsoom Baboyan 40054395

Thomas Gauvin 40064853

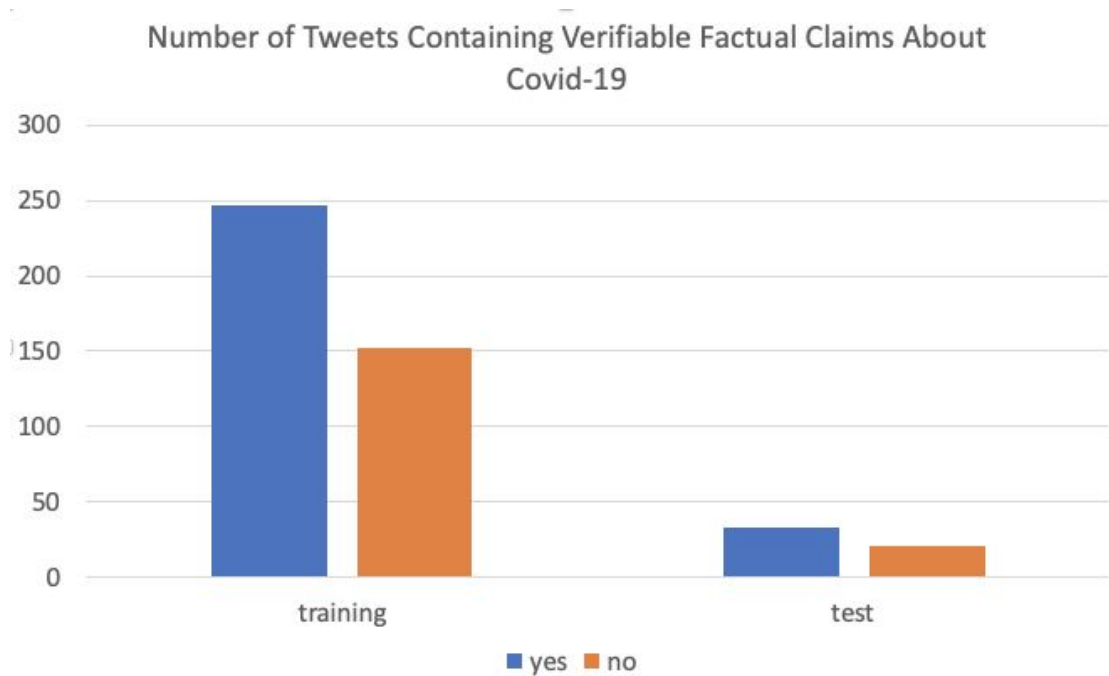
Zohal Mir 40033246

Analysis of initial datasets

Observations

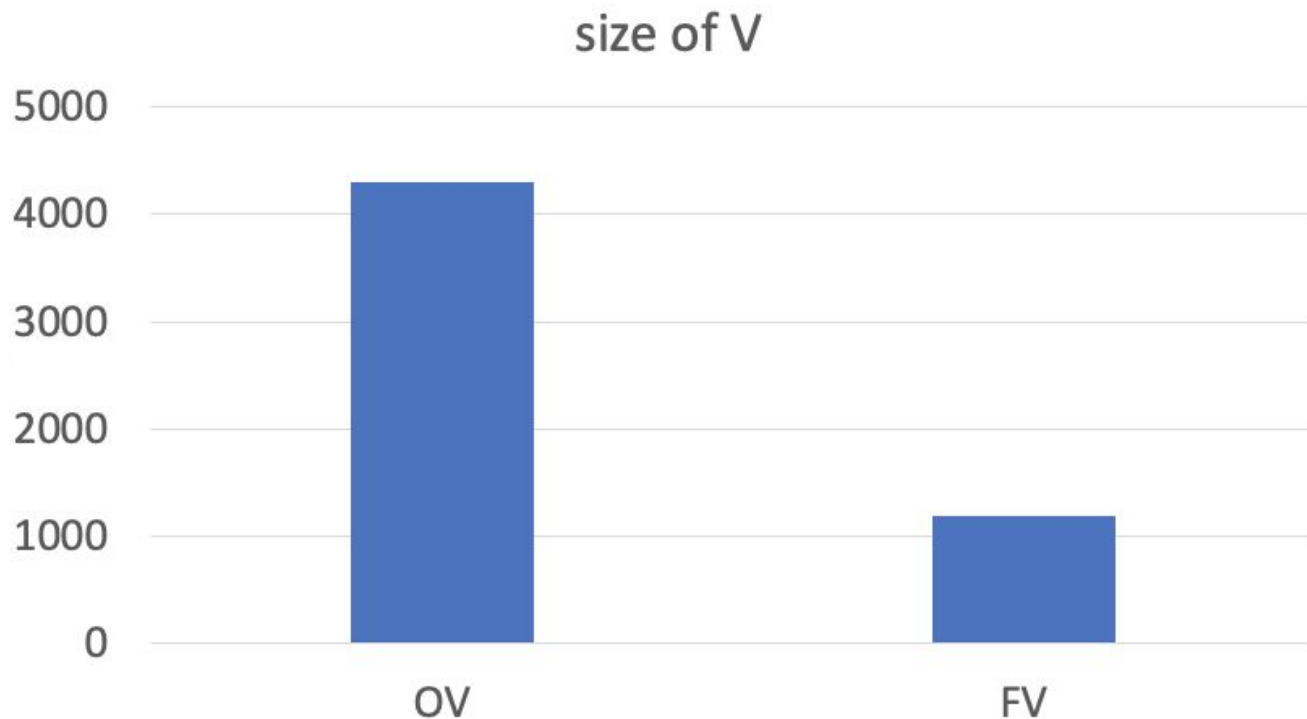


- The dataset is unbalanced.
 - Biased towards Yes.



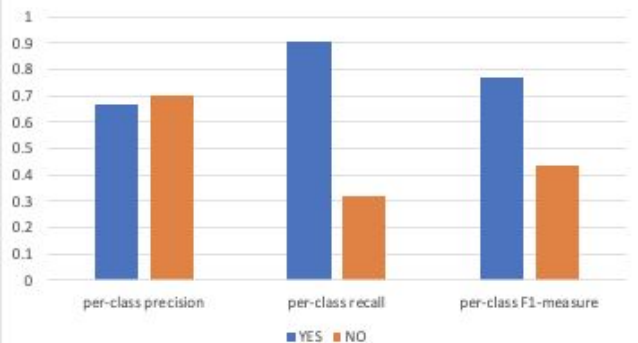
Comparison of Models

Comparison of Vocabulary Size

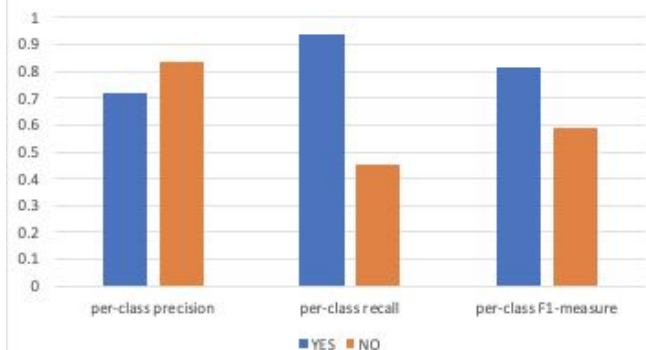


Models' Performance Comparison

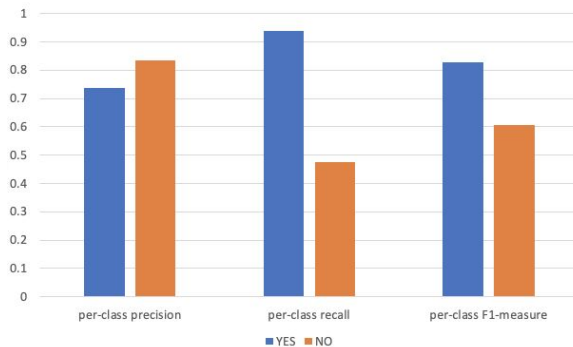
Naive Bayes Classifier with Original Vocabulary



Naive Bayes Classifier with Filtered Vocabulary



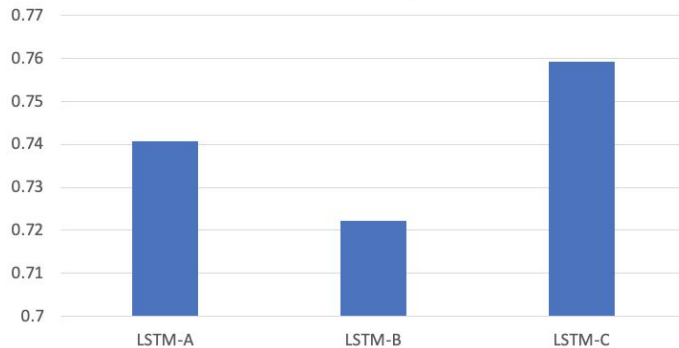
LSTM Classifier 10th Run



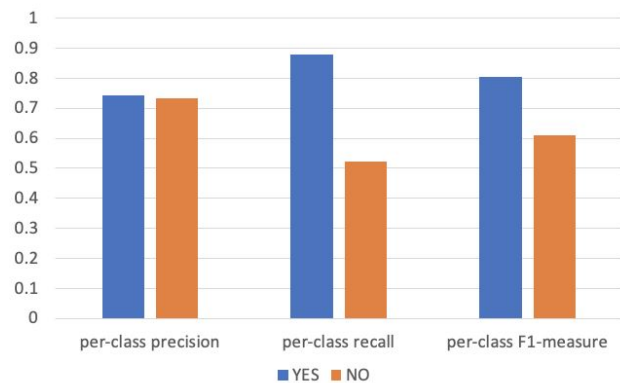
LSTM: 3 runs (A, B, C)



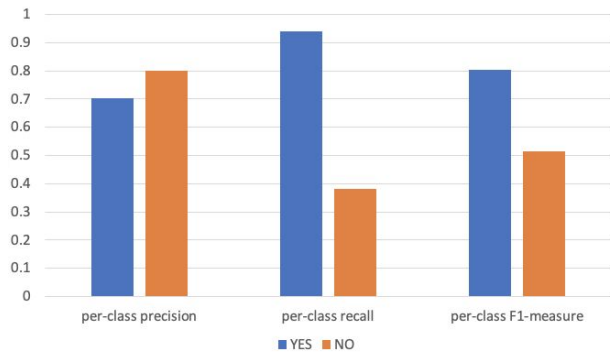
LSTM Accuracy



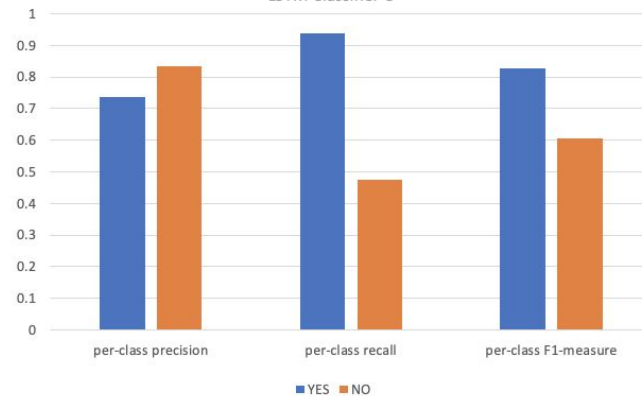
LSTM Classifier A



LSTM Classifier B



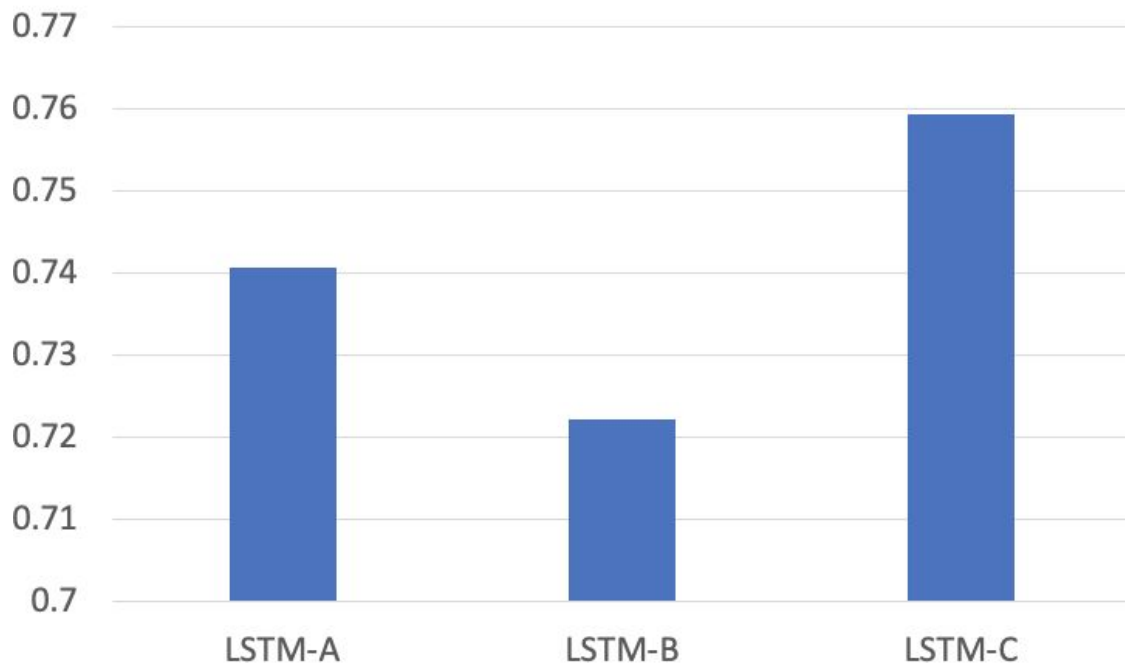
LSTM Classifier C



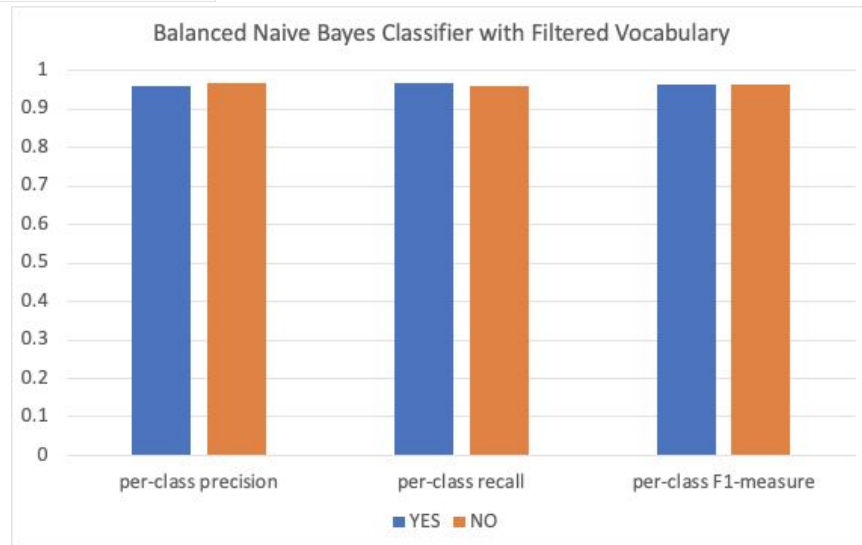
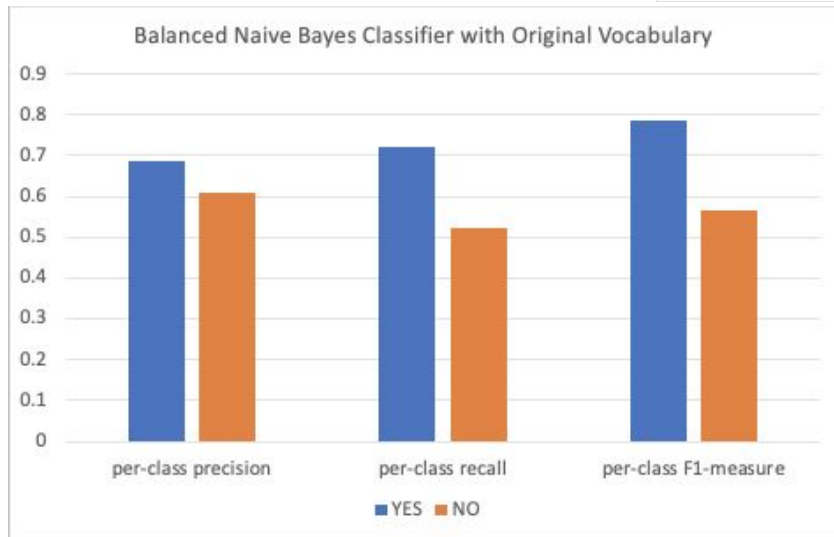
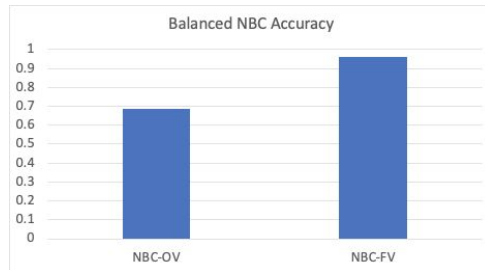
Comparison of Accuracy



accuracy



Models' Performance with Balanced Dataset



Each Member's Contribution

Each Member's Contribution



- **Hambrsoom:**
 - Code refactoring and fixing code issues
 - Perform analysis and made the powerpoint
- **Thomas:**
 - Implemented the code base for the assignment
- **Zohal:**
 - Code refactoring and fixing code issues
 - Perform analysis and made the powerpoint
- **Everyone:**
 - Pair programming
 - Review each other's work
 - Put everything together



Thank you!
Any Questions?