

# Applying Deep Learning Models on Social Media Sentiment Data in order to Improve Airline Customer Experience

Jason Mullen   Mei Zhong   Chiemela Amaefule  
College of Natural Sciences & Mathematics, University of Houston

## Abstract

As the Internet and social media becoming widespread, people can not only express their feelings and sentiments on social media, but also these sentiments and feedback can affect industries and companies' decision. This study analyzed how airline travelers expressed their feelings and sentiments on Twitter in February 2015 towards six airline companies, and figured out a method for the airlines to improve their customer services and experience. This study used Word2Vec deep learning model to convert text into vector, and used RNN, NLP and LSTM to build our models. By comparing the accuracy between different models, this study found that better performance was achieved by LSTM model. This project also showed that how social media public opinion can improve the private sector's products, services and technology.

## Background

Nowadays, reviews, surveys and tweets have become popular ways to customers to express their opinions and companies to collect feedbacks. Meanwhile, sentiment analysis allows researchers to detect positive or negative feeling in a text, and it rapidly becomes an essential tool to monitor and understand sentiment in all type of data.

In this study, the researchers decided to use the airline travelers' tweets and built RNN and various LSTM model to perform the sentiment analysis.

By analyzing the sentiment feedback and the reasons for these sentiments, this study will be able to help the airlines improve their services and customer experience, as well as further understanding how the social media affect the quality of good and services we receive in the daily life.

## Methods

- 1.Word2Vec:** With this technique, the researchers cleaned the tweets by taking out the stopwords, @ symbols and punctuations. It also made all the words lowercase, converted them to `t o k e n s`, then to vectors and applied the model.
- 2.RNN:** The input is expected to be a sequence of words or even paragraphs of words. The output can be a regression output with continuous values which represent the likelihood of having a positive sentiment.
- 3.LSTM:** LSTM has long term memory compared to RNN which can only learn about recent information. LSTM can help solve this problem as it can understand context along with recent dependency. Since this research is dealing with tweets, so it's necessary to memorize the previous words and understood the context in order to get the sentiment. Therefore, LSTM's ability to understand context is very useful.
- 4.Stacked LSTM:** This project used stacked LSTM in order to improve upon the regular LSTM Model.

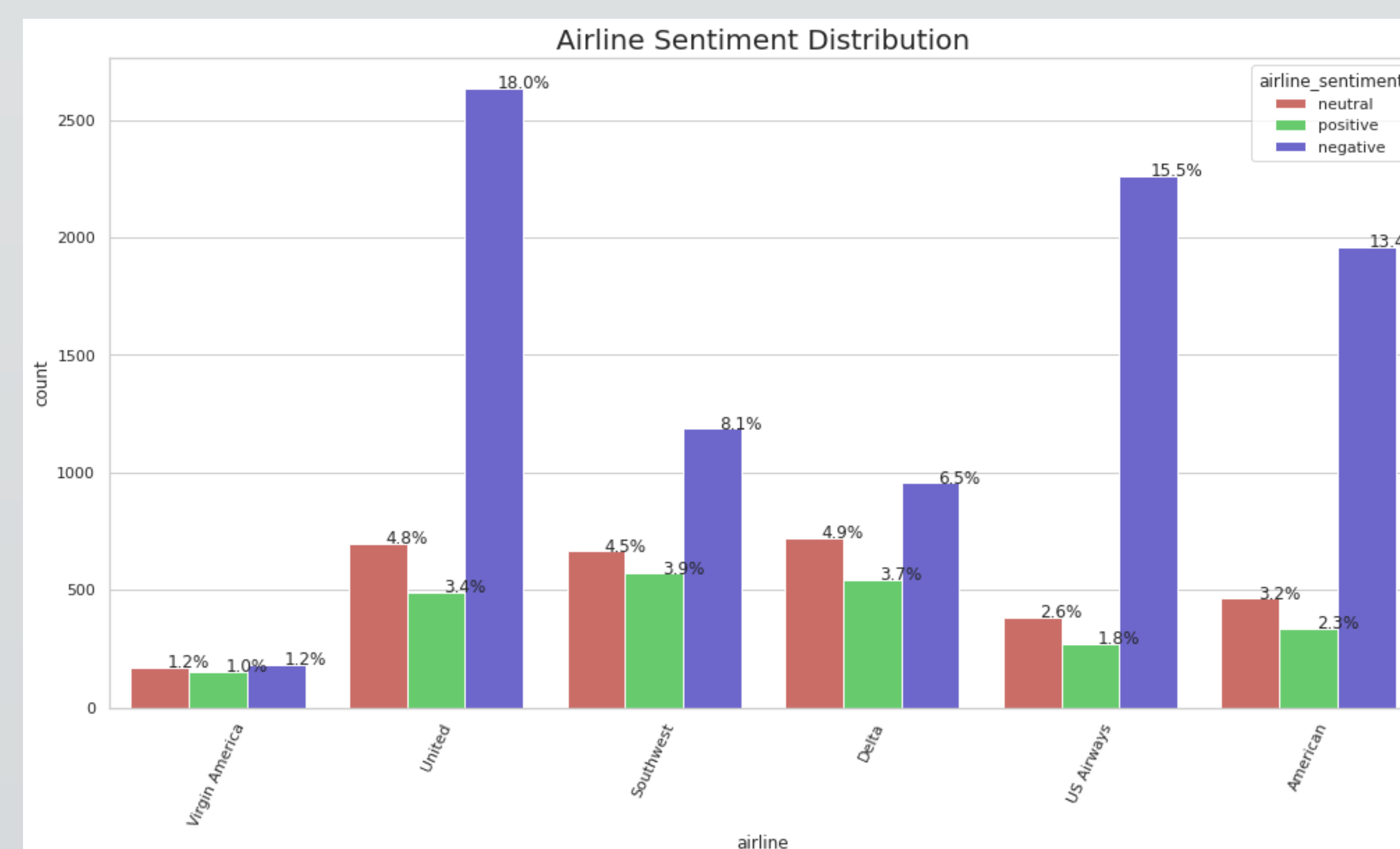


Fig 1. Airline Sentiment Distribution of six Airplane Companies

## Results

- RNN: 73.4%
- LSTM: 73.7%
- Stacked LSTM: 72.6%

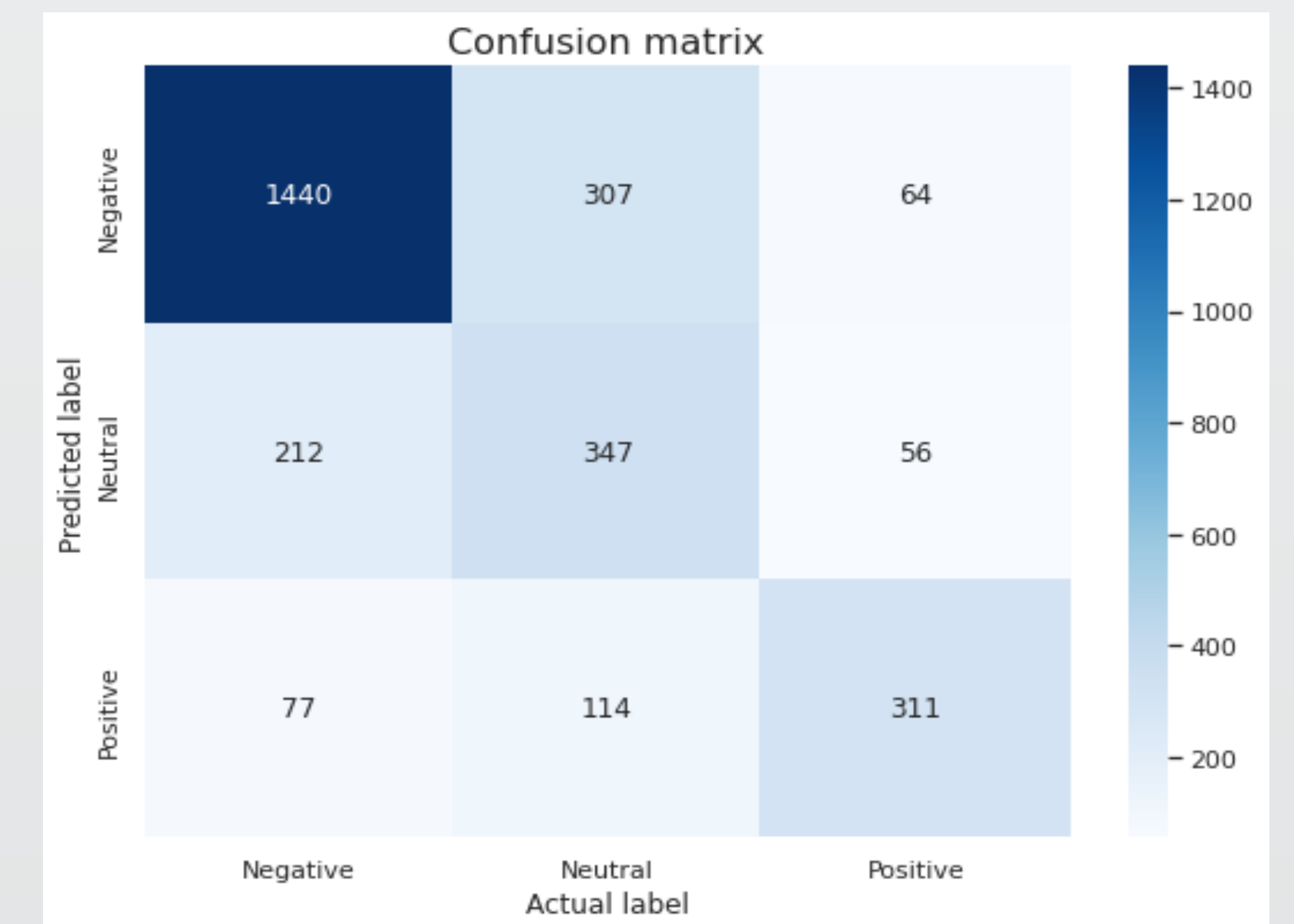


Fig 2. Confusion Matrix of LSTM Model

## Conclusion

Other studies conducted on sentiment analysis used RNN, CNN, LSTM, and Stacked LSTM. Their highest accuracy was achieved by their LSTM model with one report reaching a 91% accuracy with stacked LSTM.

This project built RNN, LSTM, Stacked LSTM and achieved different results. It achieved a higher result with LSTM model.

## Future Direction

- In the future, this project can be improved by using other classification tools such as Support Vector Machines and K-Nearest Neighbors as well as CNN and comparing the results.
- This model can also be used to conduct analysis on students feelings towards the college experience to provide feedback to faculty and staff.
- Real time sentiment analysis.

## Acknowledgments

We would like to thank Dr. Nouhad Rizk from the Department of Natural Sciences and Mathematics. This work was made possible through the Hewlett Packard Enterprise Data Science Institute at the University of Houston.