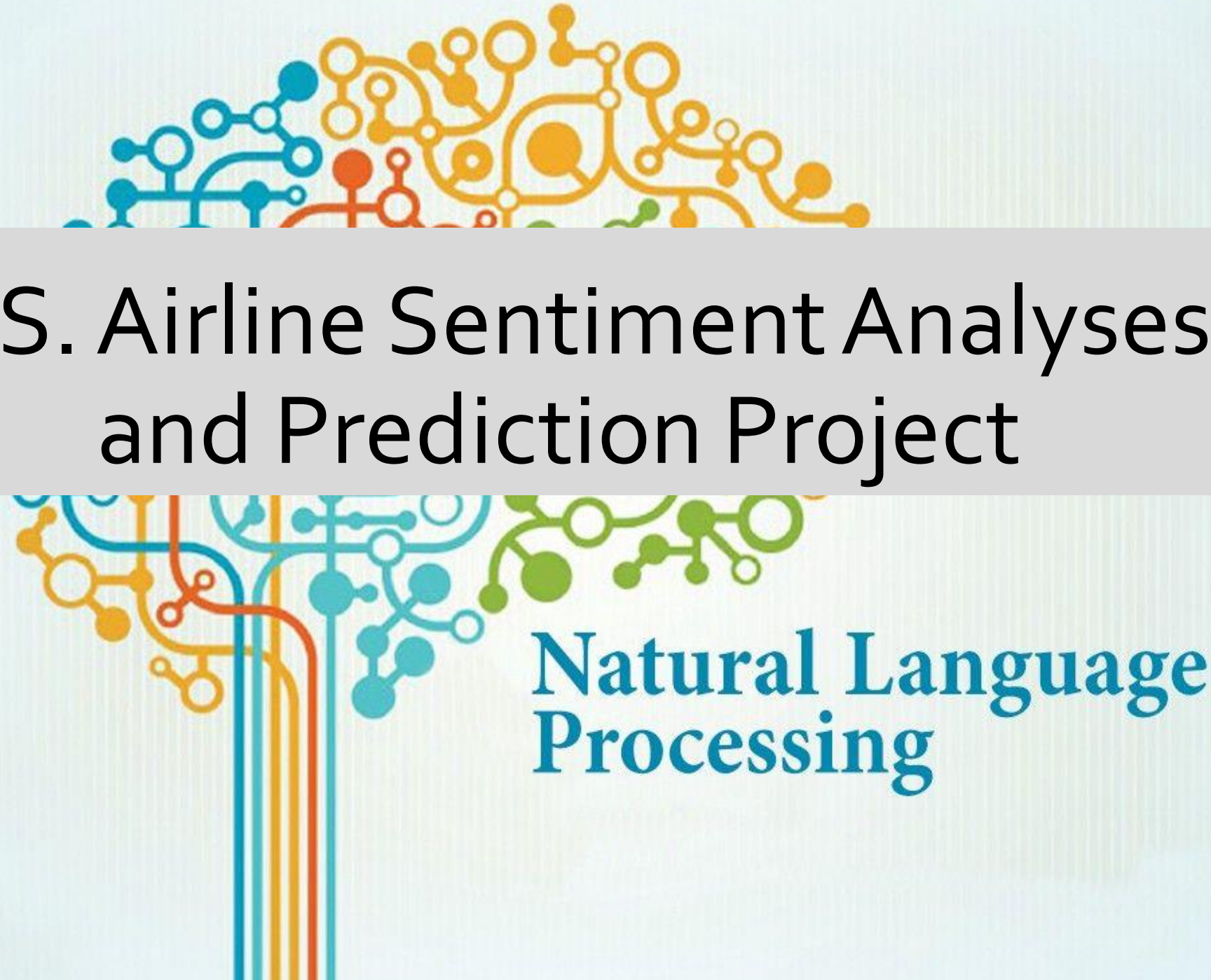
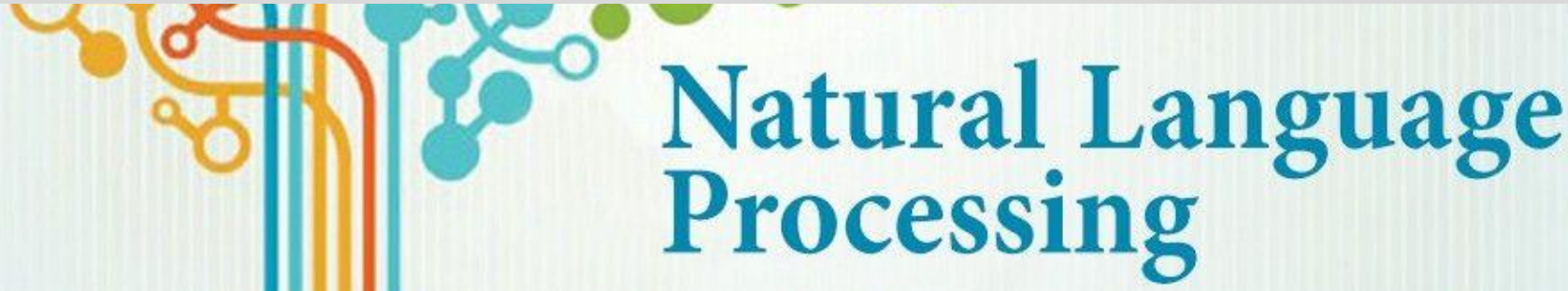
U.S. Airline Sentiment Analyses and Prediction Project



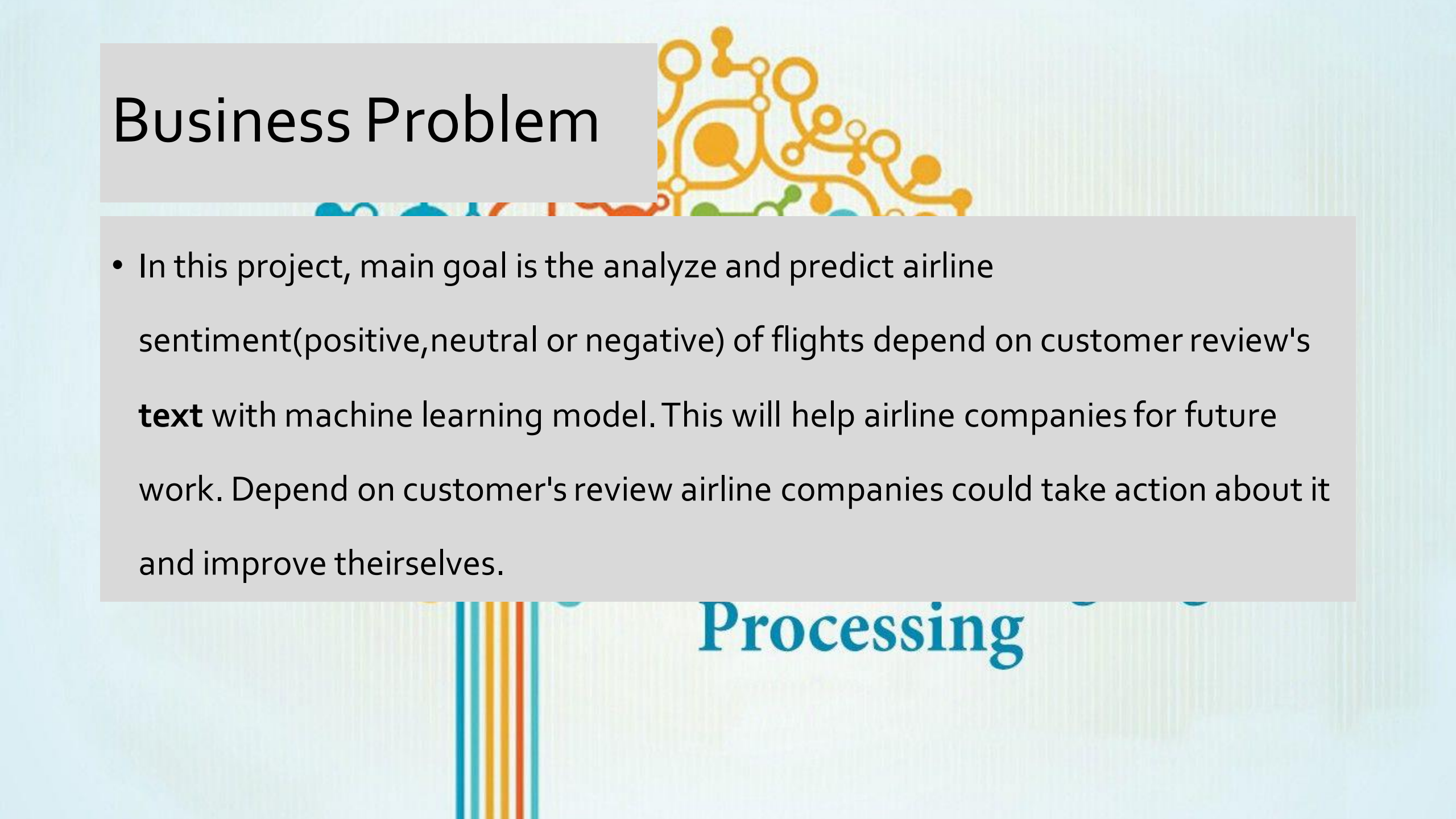
Business Value

- There are six different airline companies in this dataset and their customers still complaining about some problems with their services/flights. We are going to analyze and making machine learning project for how airline companies could improve theirselves with our findings.



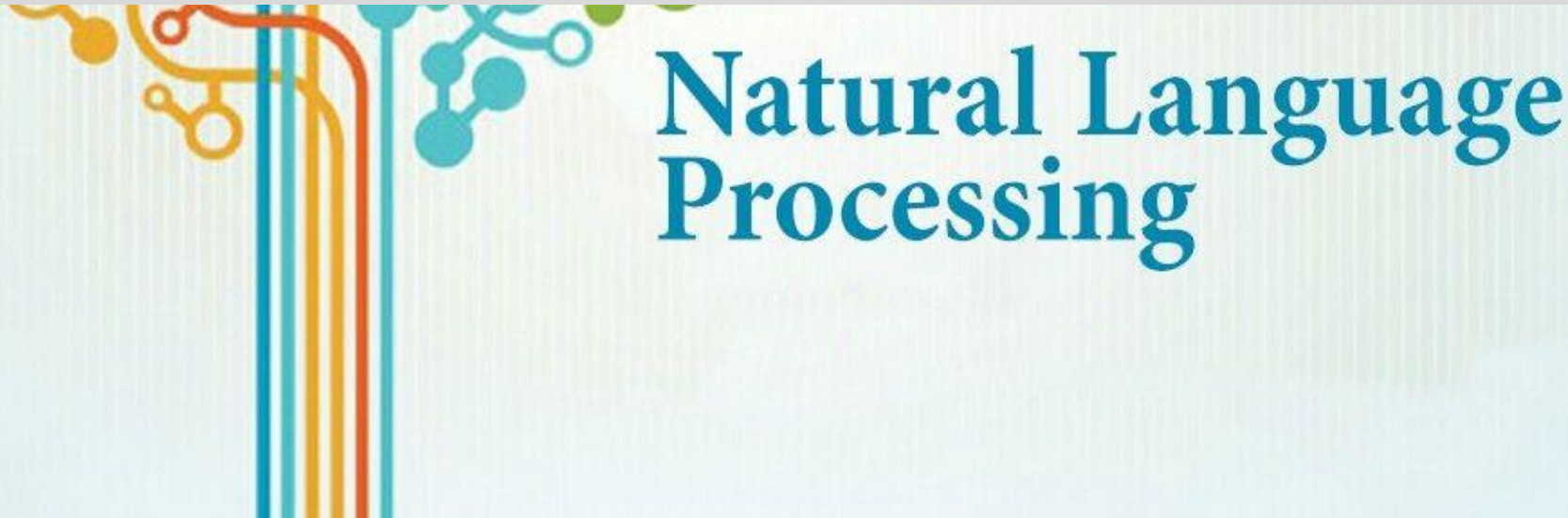
Business Problem

- In this project, main goal is the analyze and predict airline sentiment(positive,neutral or negative) of flights depend on customer review's **text** with machine learning model. This will help airline companies for future work. Depend on customer's review airline companies could take action about it and improve theirselves.

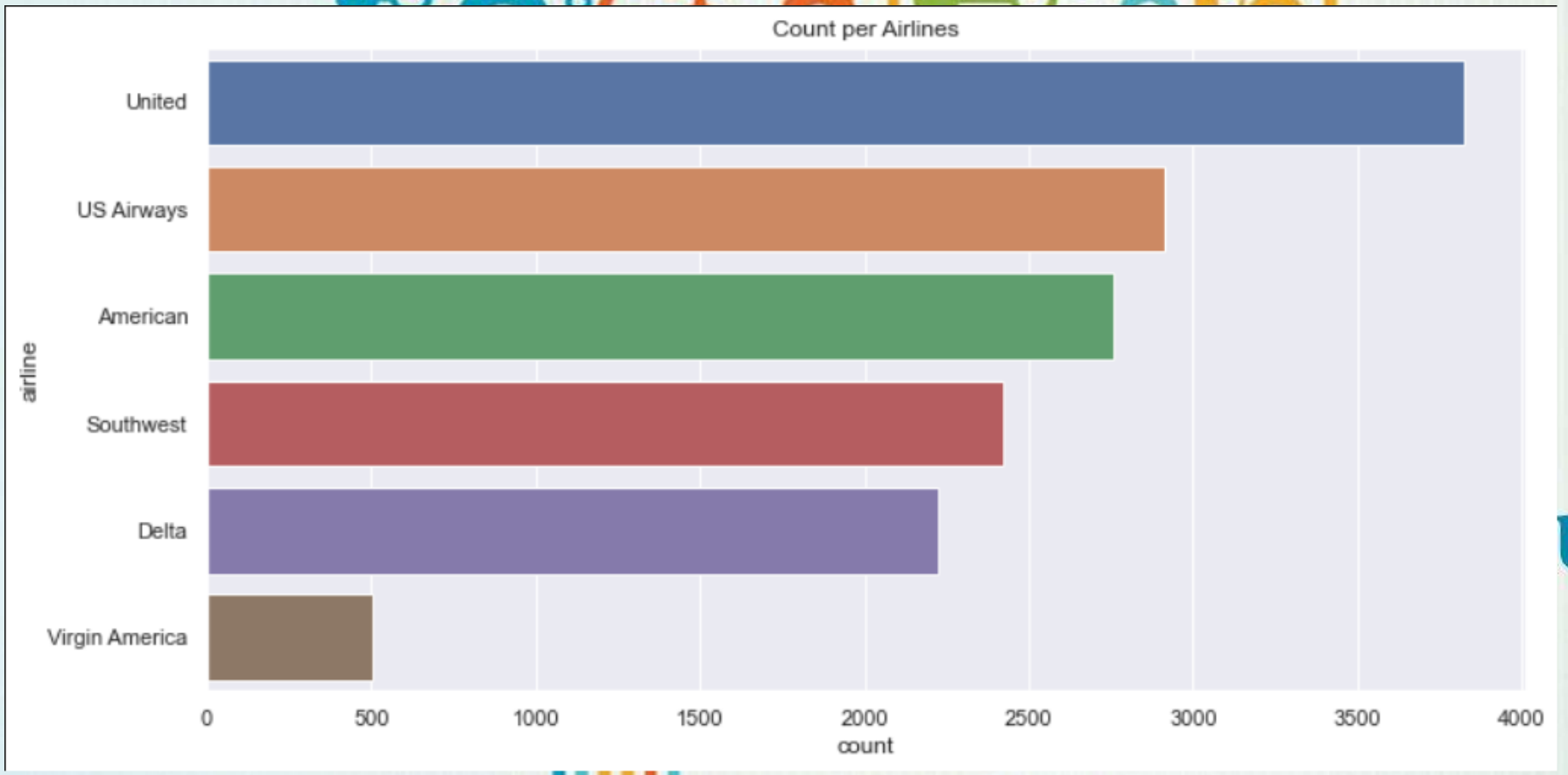


Data Understanding

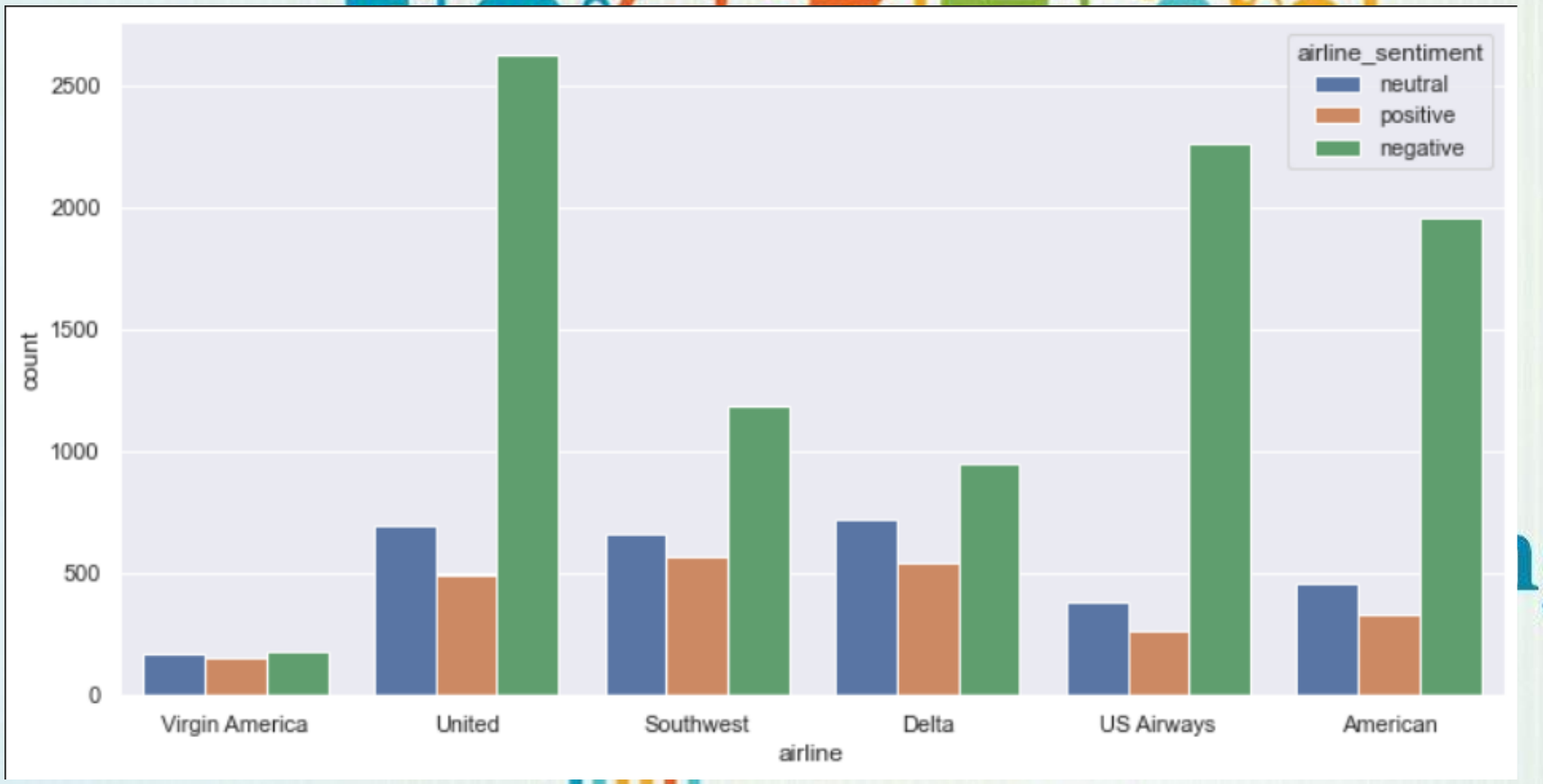
1. We have 6 different major U.S. airline companies; United, Us Airways, American , Soutwest, Delta and Virgin America.
2. All the tweets scraped at 2015 February.
3. We have three different target category as positive, neutral or negative depend on tweet.



Tweet Counts per Airline Companies

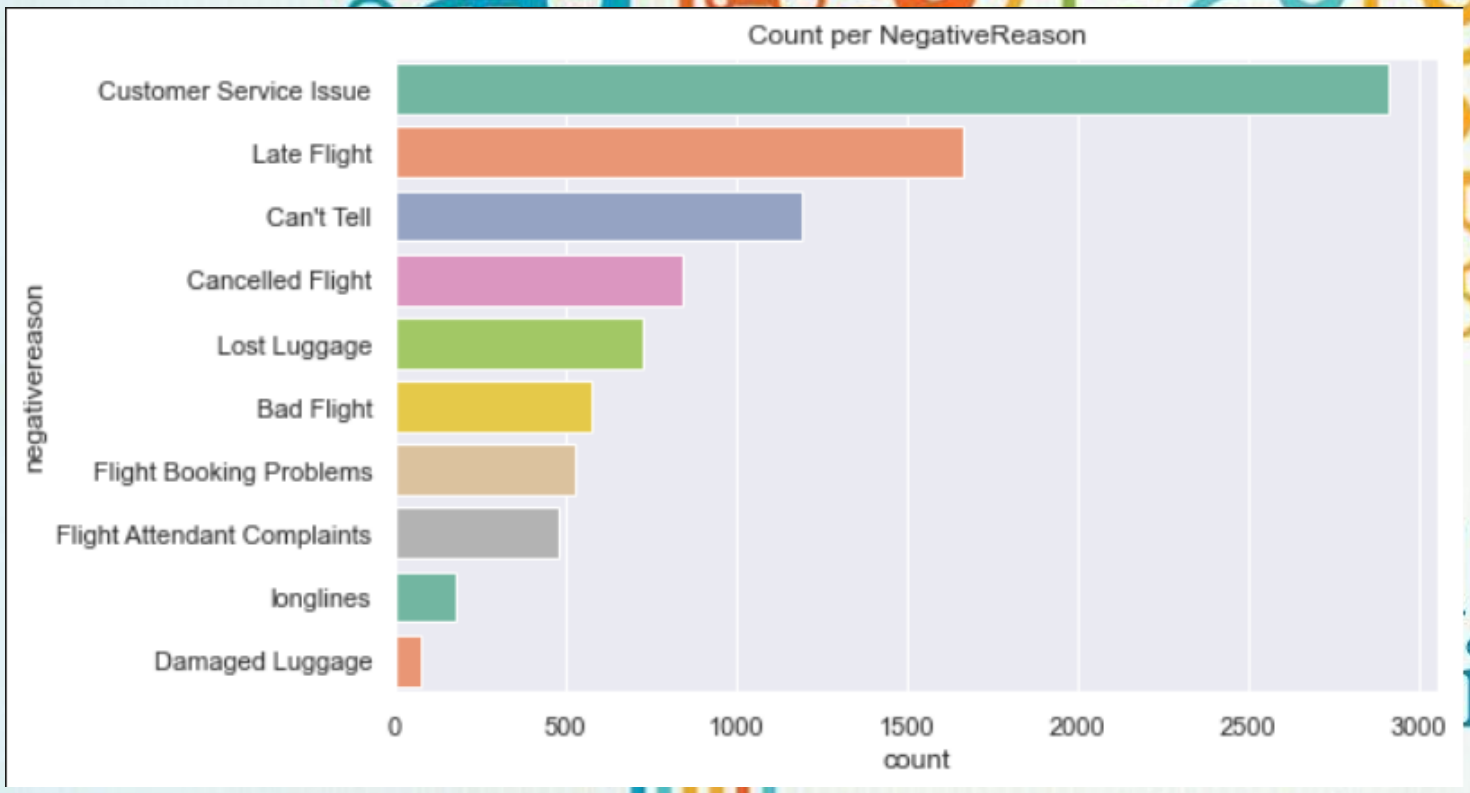


Sentiment per Airline Companies



In this graph, we can see with greens are negative tweets and most of them are negative. Specially United, Us Airways and American Airlines have the most negative customer reviews.

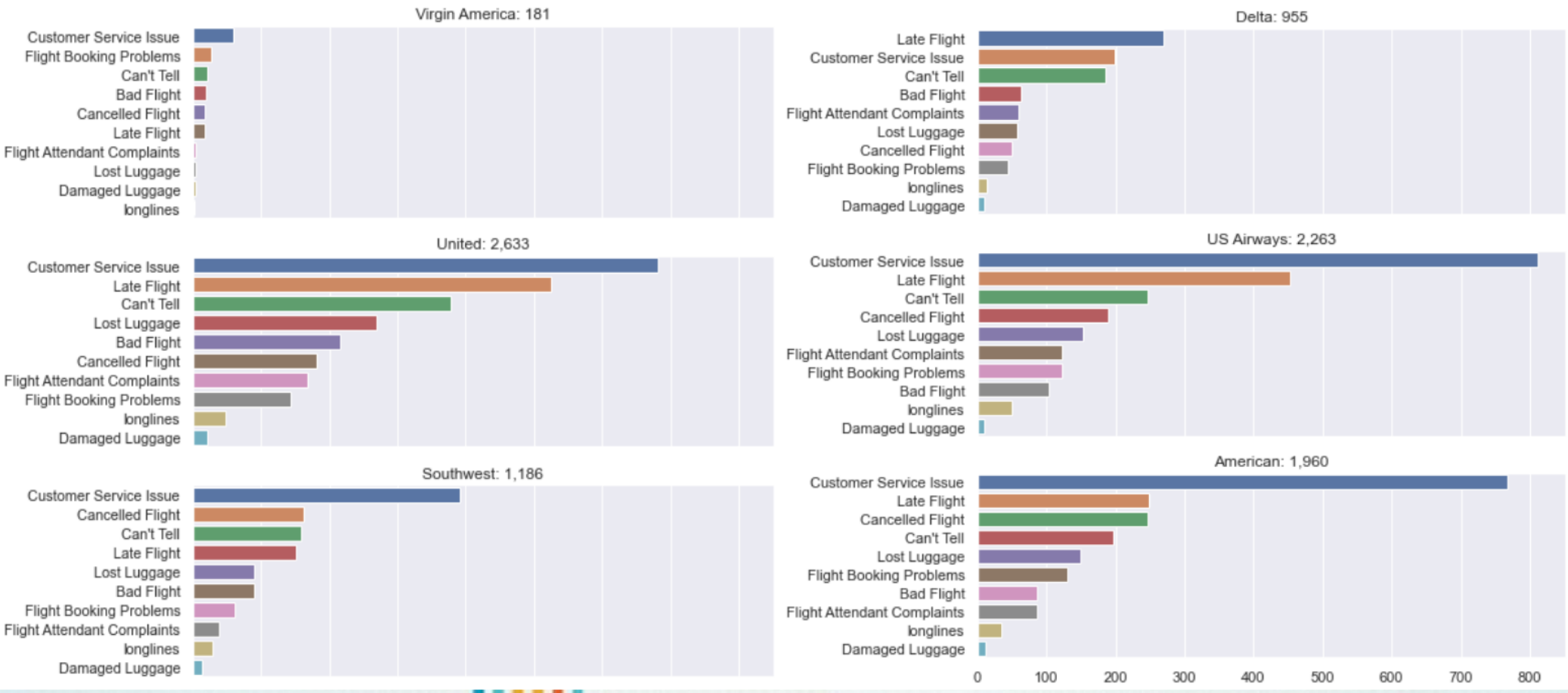
Negative Reasons



This visual shows us general negative reasons from customers. Customer service issue and late flight leads for negative reasons.

Language
ing

Negative Reasons per Airline Companies

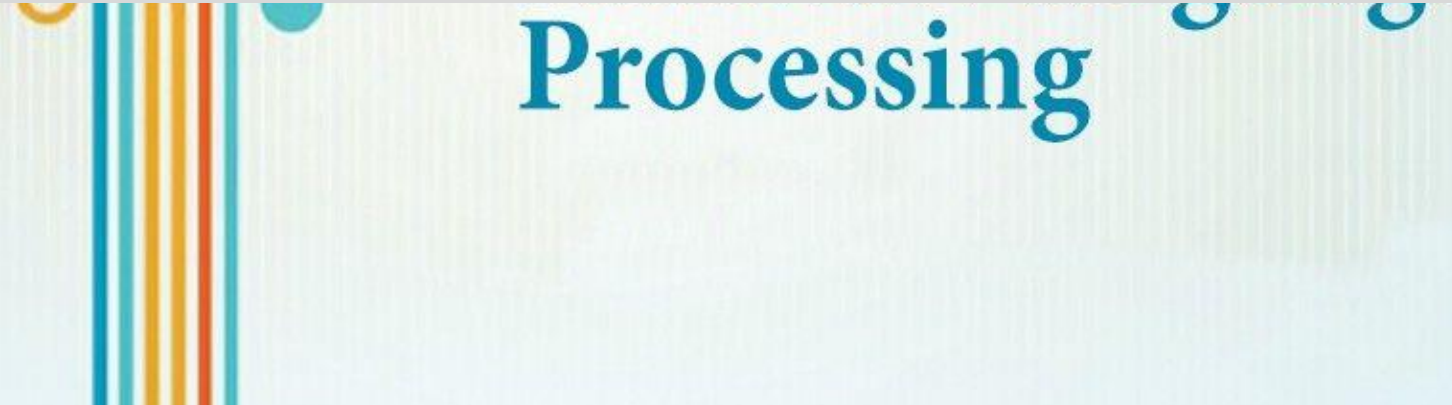


Conclusion

In this project tried different machine learning models to find best result. Some of basic models name are ; Logistic Regression, Random Forest, Gradient Boosting. But improved model called neural network model gives us best result as **92%** accuracy on dataset.

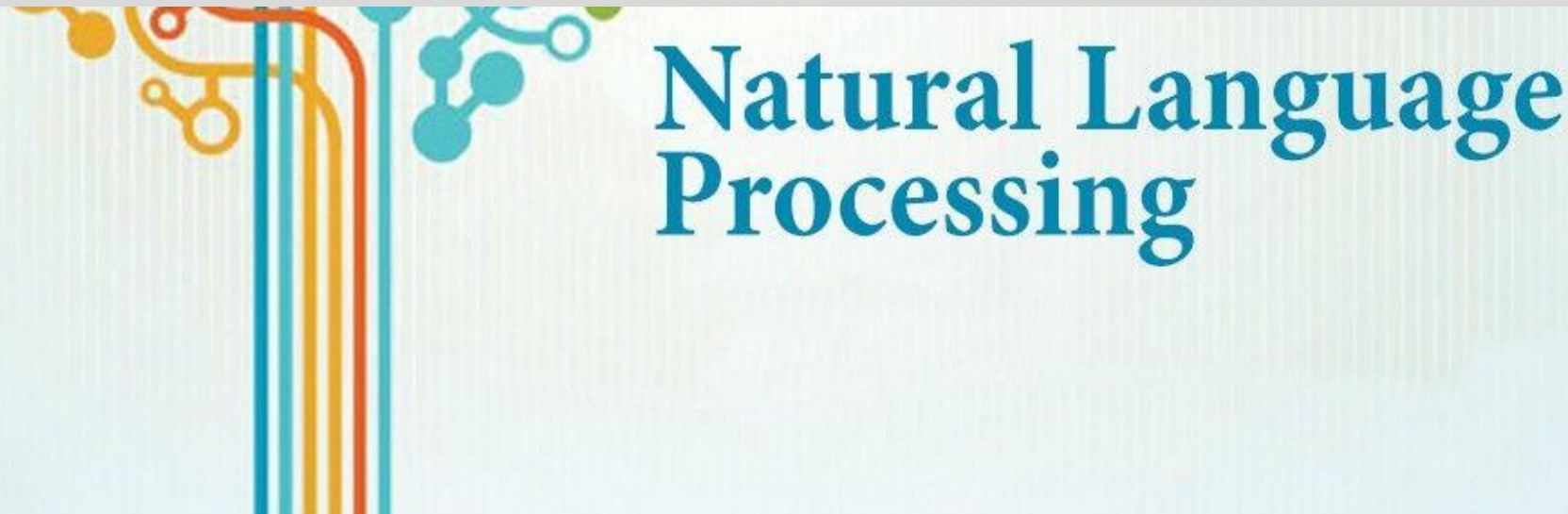
What this number mean is model will predict 92% true (positive,neutral or negative) on given tweet. For example;

We will give our model as tweet "didn't enjoy flight". Model result will be 'Negative'.



Future Step

- We can work on detail text and improve our model.
- For example we still get little bit of false positive,neutral or negatives. We would work why is that and what we could do more on our model.
- Also gather more tweets about these airline companies.



By Huseyin Caglar

huseyincaglar@outlook.com