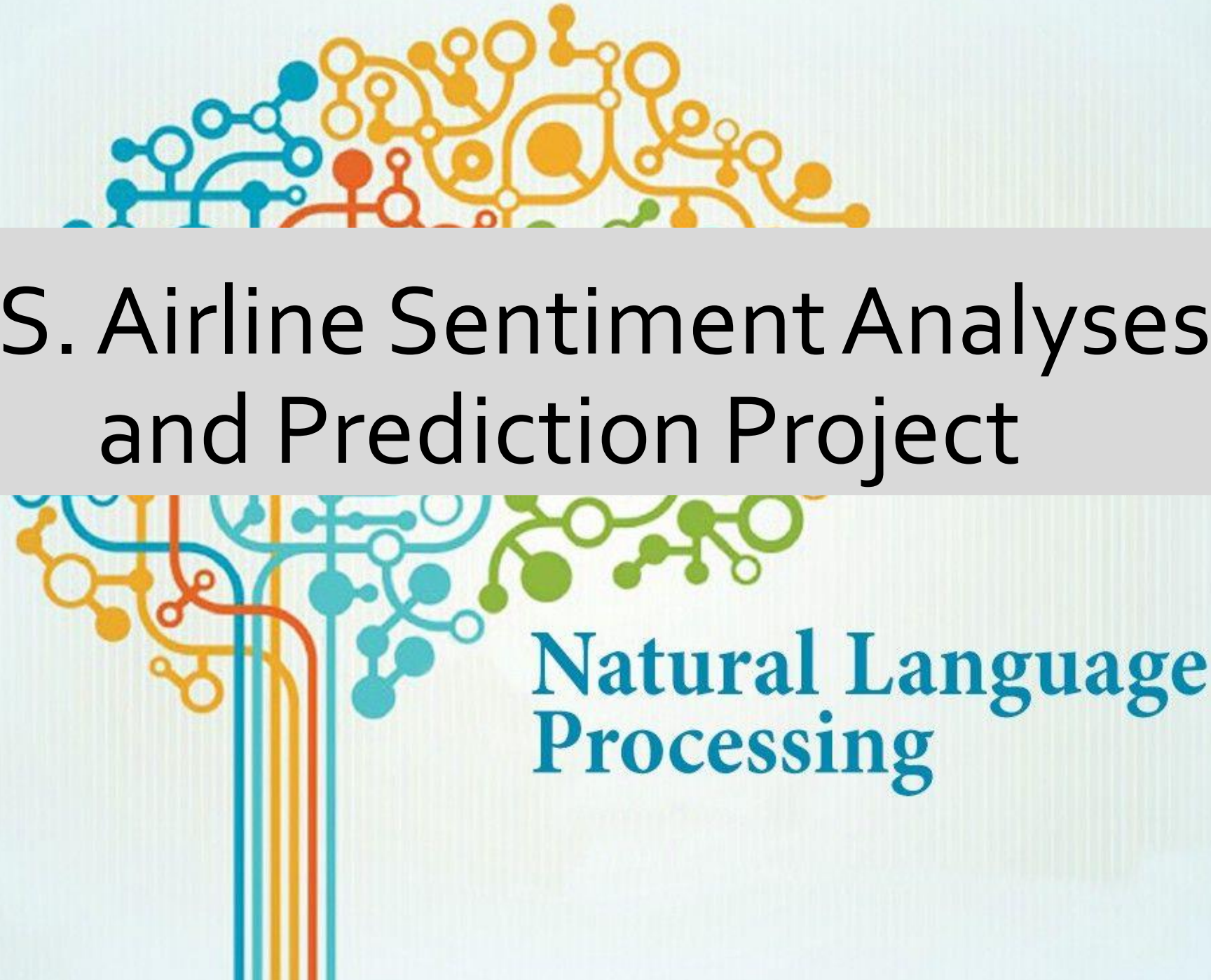
U.S. Airline Sentiment Analyses and Prediction Project



Business Value

There are six different airline companies in this dataset; United, US Airways, American, Southwest, Delta and Virgin America. And their customers still complaining about some problems with their services/flights. For an airline company one customer, customer's review, one cancellation flight, one hour or one minute sometimes seconds too much important for a business value. Because of business reputation and business economic status. Instead of other industries economic status is more important at airline industry because this is a transportation company and losing every second for every mile flight. We are going to analyze and making machine learning project for how airline companies could improve ourselves with our findings.

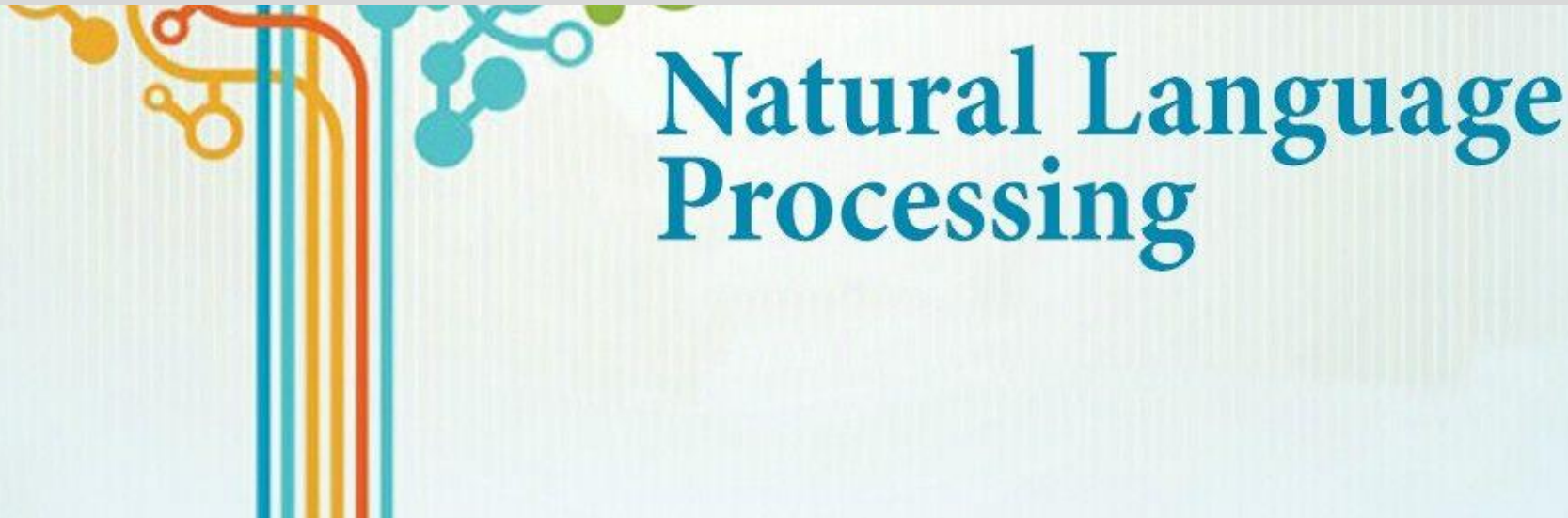
Business Problem

- In this project, main goal is the predict airline sentiment of flights with machine learning model. Our problem is customers satisfaction of flights.Some customers not only half satisfied, almost completely not satisfied and have some problems like; customer service issue, late flight, cancellation of flight etc. This problems will make specific airline company to lose money.Since every seconds important for an airline company, we are going to analyze why is that and making machine learning model to prevent at the future. Depend on customer's review(positive , neutral or negative) airline companies could take action about it.

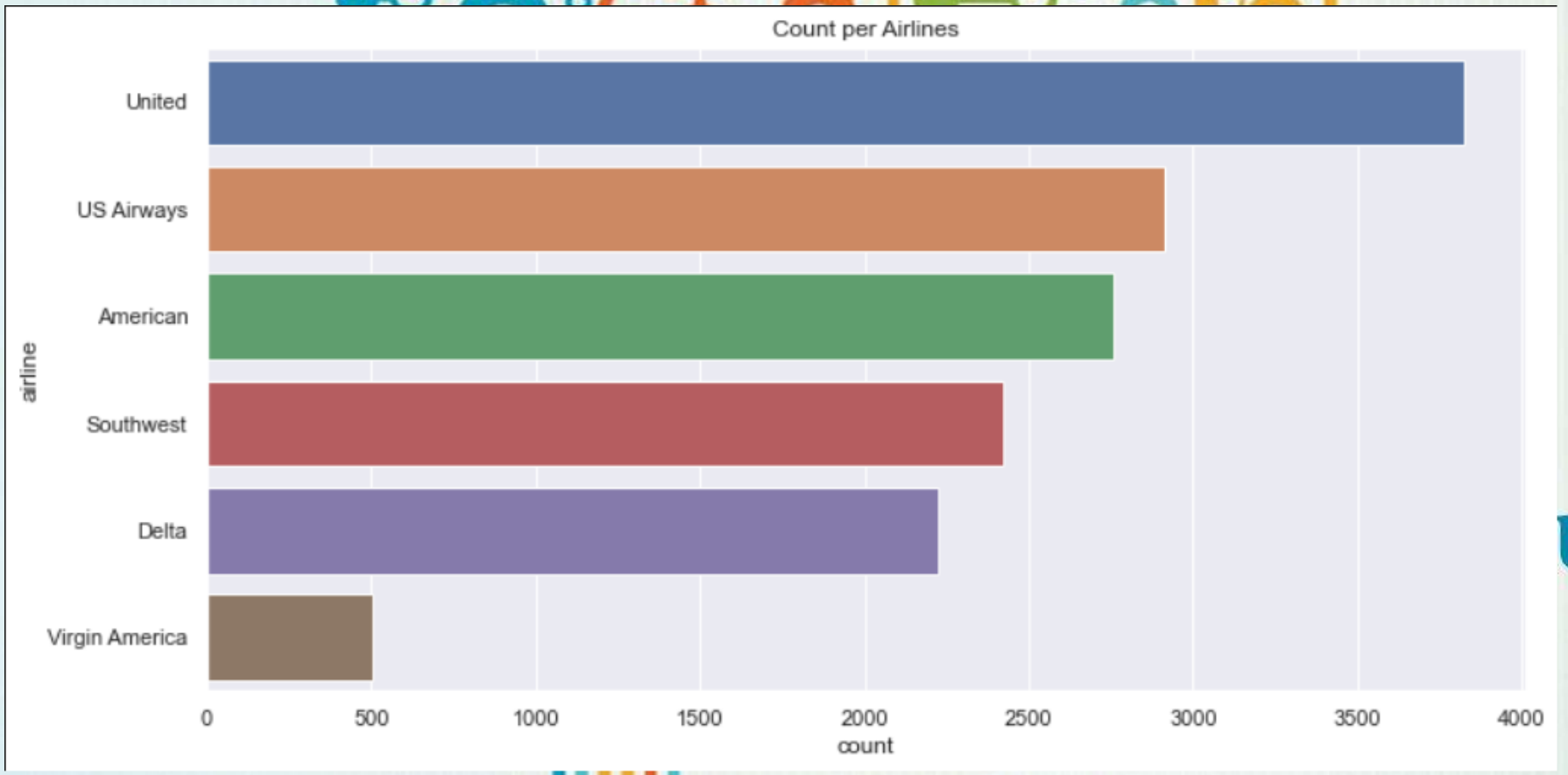
Processing

Data Understanding

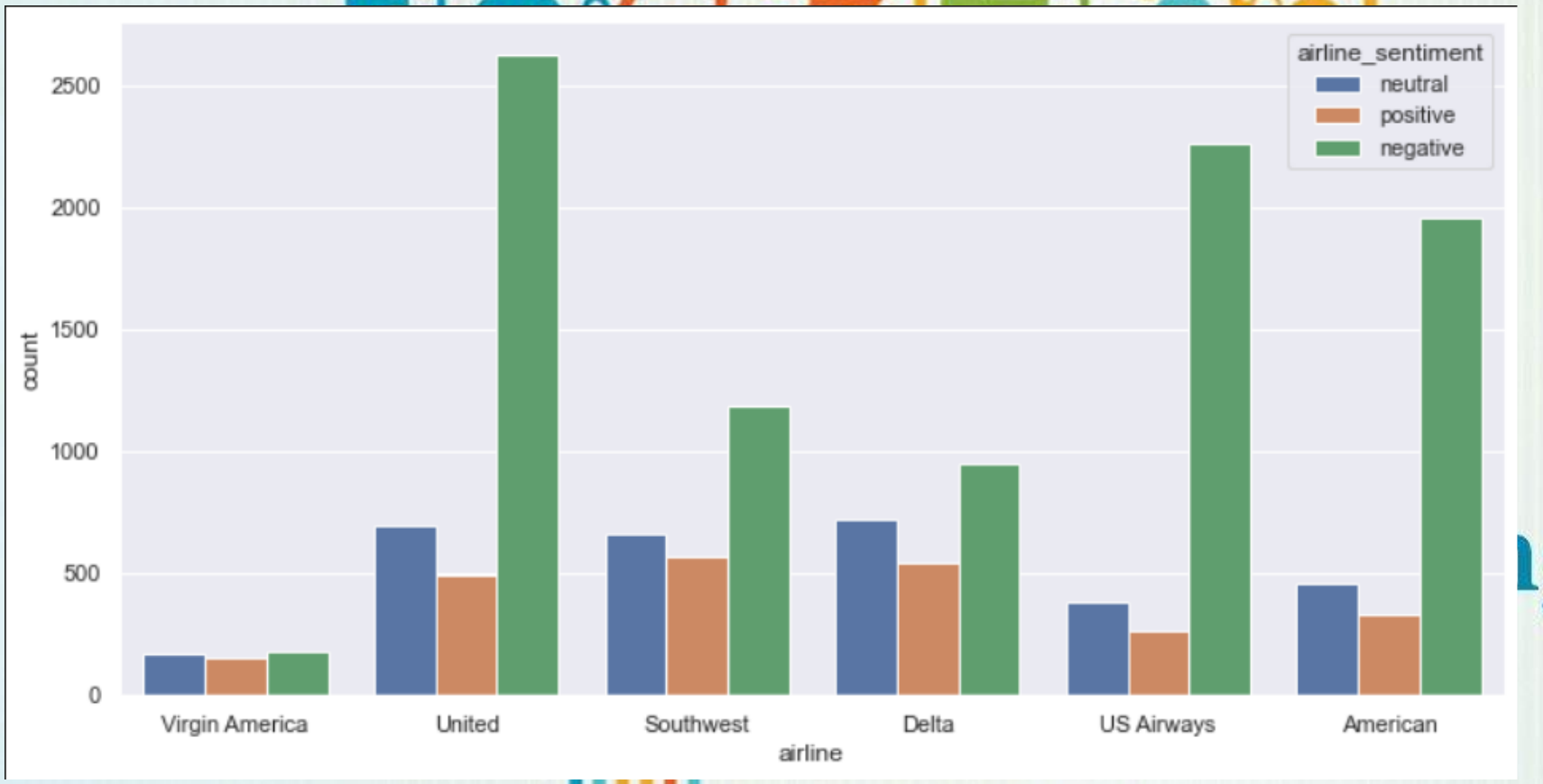
1. We have 6 different major U.S. airline companies; United, Us Airways, American , Soutwest, Delta and Virgin America.
2. All the tweets scraped at 2015 February.
3. We have three different target category as positive, neutral or negative depend on tweet.



Tweet Counts per Airline Companies

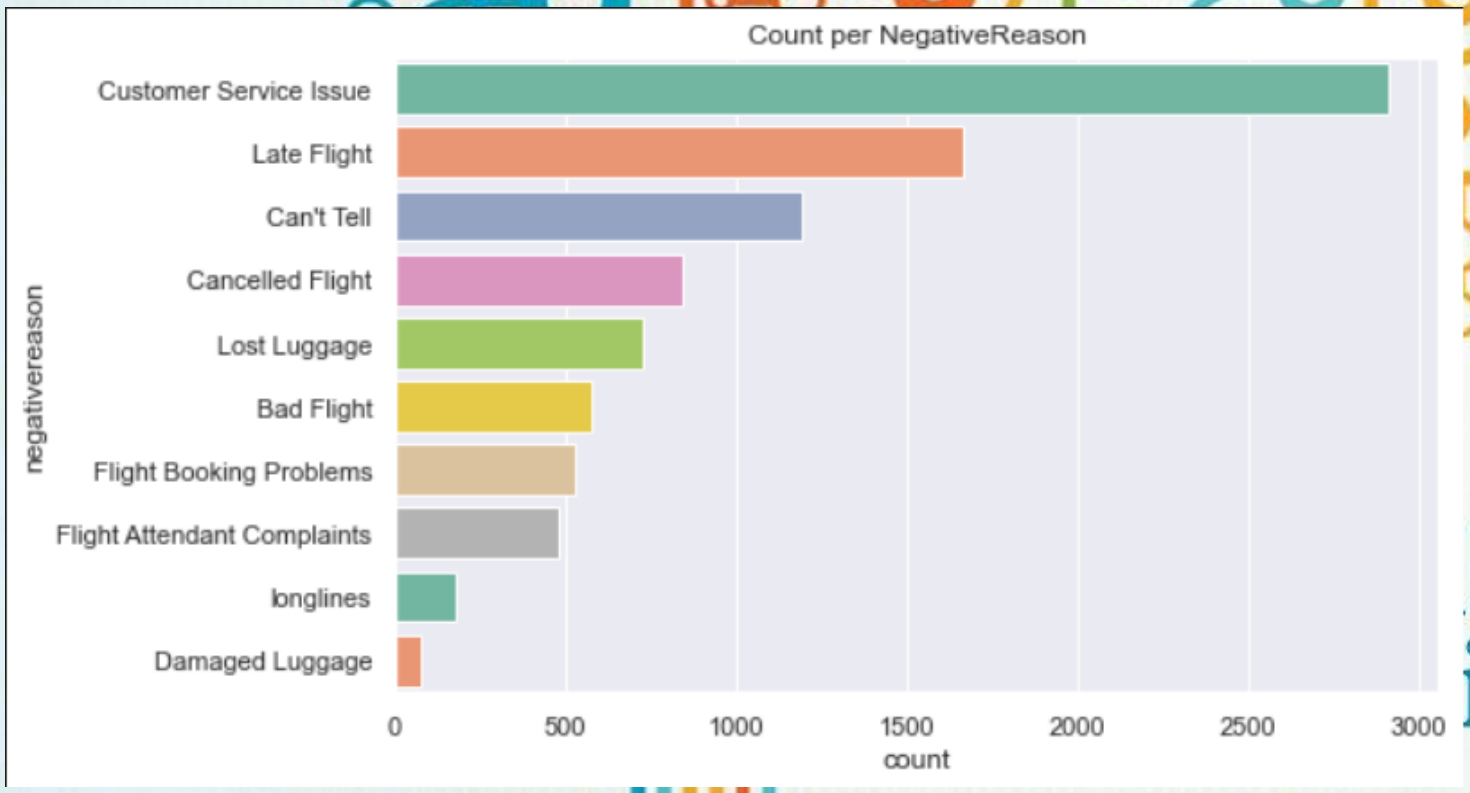


Sentiment per Airline Companies



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.

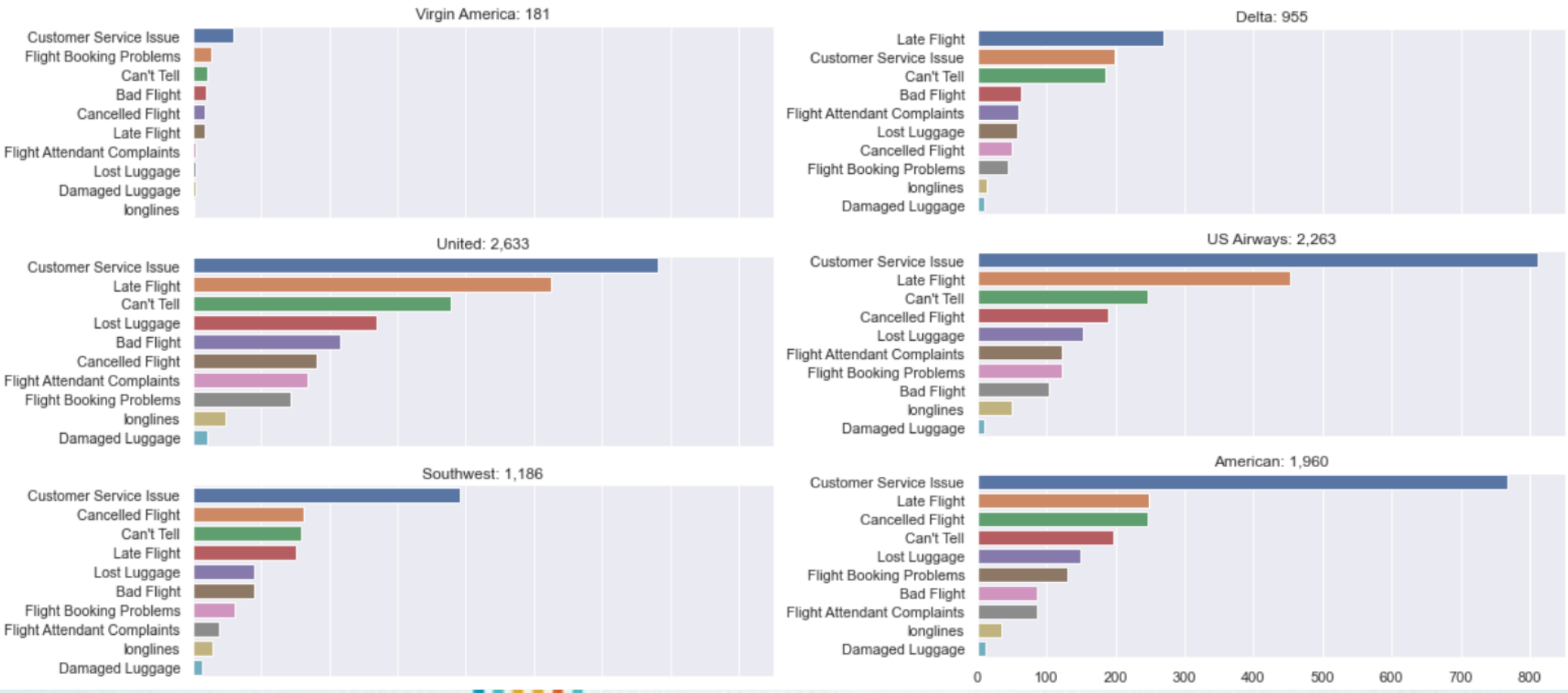
Negative Reasons



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



Recommendations

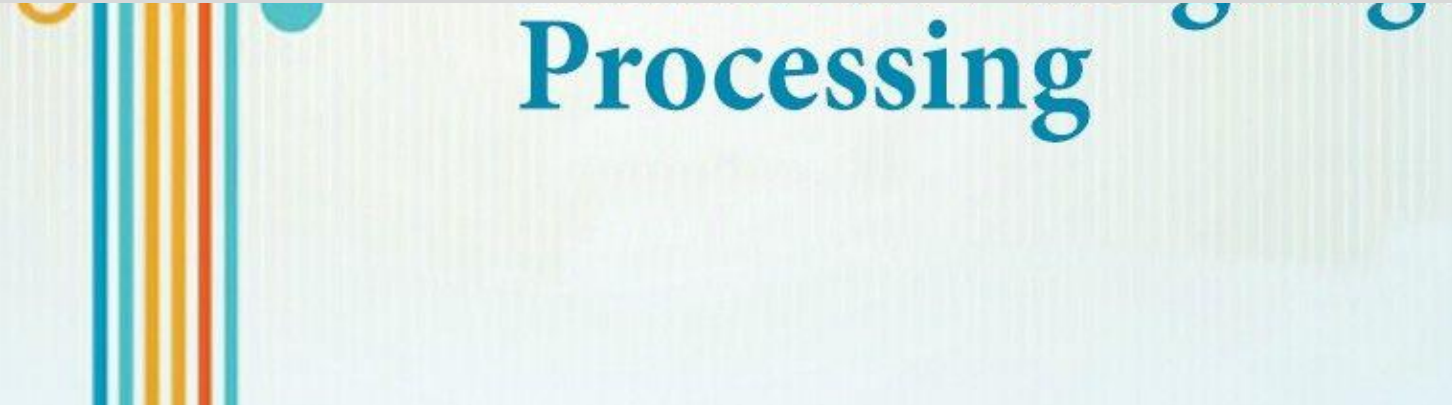
- For all the 6 companies should work on customer issue problems.
- Virgin America needs to check booking problems.
- Delta Airlines needs to look into late flight problems.
- Overall late flight and cancellation flight problems needs to be solve for better service/more satisfied customer.

Conclusion

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

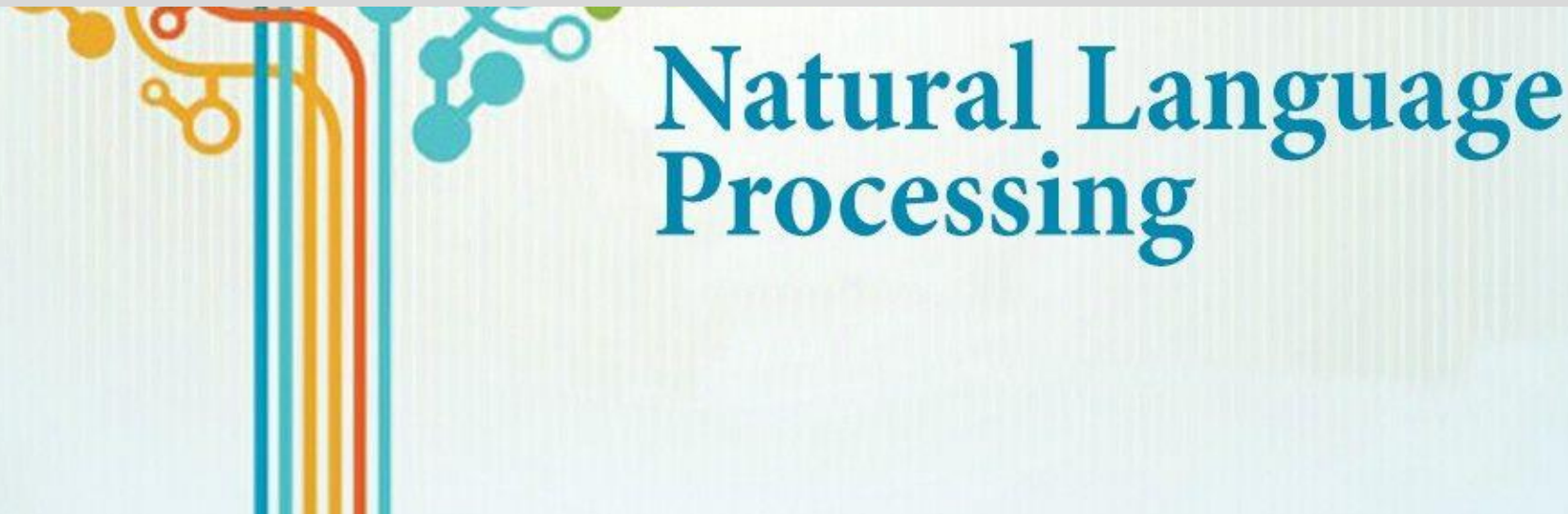
What this number mean is model will predict 92% true (positive,neutral or negative) on given tweet. For example;

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