

**TWITTER US AIRLINE SENTIMENT**

Course Instructor: **Prof. Savita Sehrawat**

Ravinderjit kaur (0763523)

**TWITTER US AIRLINE SENTIMENT**

**Table of Contents:**

**INTRODUCTION………………………………………………………………………3-5**

* Abstract………………………………………………………………………………3
* Research Questions………………………………………………………………….4
* Related Work……………………………………………………………………….4-5

**LITERATURE SURVEY……………………………………………………………………5**

**DATA COLLECTION…………………………………………………………………5-14**

* EDA (Exploratory Data Analysis) ………………………………………………7
* Data Visualization……………………………………………………………8-14
* Visualization Of Graphs………………………………………………8-13
* Word Clouds……………………………………………………………13-14

**MATERIALS AND METHODOLIGIES……………………………………………15-26**

* Complete pre-processing
* Partial pre-processing
* TF-IDF approach
* Feature-Extraction Methods
* Random Forest Classifier
* Naïve Bayes
* SVM
* Decision Trees
* XG-Boost
* Evaluation of parameters

**CONCLUSION……………………………………………………………………………27**

**REFERENCES…………………………………………………………………………28**

**A****bstract:**

In the past several years, there has been a surge in the utilisation of data from social networks like Twitter to enhance political campaigns, the calibre of goods and services, sentiment analysis, etc. For many firms, classifying tweets based on user attitudes is a crucial and collaborative effort. Based on the sentiments they contain, tweets were divided into three categories: positive, negative, and neutral. Additionally, a range of machine learning classifiers were assessed using the performance measures accuracy, precision, recall, and F1. Moreover, the performance of a deep long short-term memory network was analysed on the selected dataset. The results show that the propose performs better than that of other classiﬁers. The results demonstrate that ensemble classiﬁers achieve higher accuracy than non-ensemble classiﬁers. Experiments further proved that the performance of machine learning classiﬁers is better when TF-IDF is used as the feature extraction method. Word2vec feature extraction performs worse than TF and TF-IDF feature extraction.

**INTRODUCTION**

In this study, natural language processing sentiment analysis for the Twitter US airline dataset is carried out. The text field in the dataset was divided into positive, negative, and neutral sentiment polarity. Sentiment analysis, often known as opinion analysis, is a machine learning method that airline services now extensively use to gauge client sentiment or the general consensus regarding their offerings in social media language. The airline support staff is focused on analysing social media content from blogs, online forums, comments, and reviews. This evaluation is being abused to form opinions or measure how well their services are performing.

The input data to the classification model must be closed using classification procedures in order to train the data. For the newly learned data, these models forecast the class label categories.

**Research Questions:**

• Which flights received which kind of favourable or unfavourable tweets, and about which services, will assist to sustain or change the service in accordance with customer emotion.

• Which complaints about poor treatment on aeroplanes are most frequently tweeted by passengers.

• Determine whether planes had more favourable or bad tweets, then classify the latter according to the source.

• Does flight time have an impact on service quality?

**Related Work**

Previous studies have classified tweets for several airline firms using supervised machine learning algorithms and lexicon-based techniques. The majority of scholars have examined the technique of sentiment analysis by identifying emotions in tweets. The Nave Bayes technique for sentiment analysis was used by Dutta Das et al. (2017) to analyze airline Twitter data utilizing 200 tweets sent to Emirates and Jet Airways. They enhanced the classification model and mapped the tweets into the positive, negative, and neutral categories using R and Rapid Miner tools. They stated that with a larger sample size of tweets included in the research, the results obtained using Nave Bayes classifiers were encouraging. The SMOTE method was used by Hakh et al. (2017) to address the datasets unbalanced difficulty and analyzing a collection of tweets regarding six US-based airlines that were discovered using machine learning methods. They discovered that in order to get precise findings, feature selection and over-sampling strategies are both crucial. The sentiment classification was then used, using algorithms including AdaBoost, Decision Tree, Linear SVM, Naive Bayes, Random Forest, K- NN, and Kernel SVM. Rane and Kumar (2018) cleaned the tweets using their method and pre-processing procedures. To do a phrase-level analysis that takes the word order into account, these tweets were represented as vectors using the deep learning concept (Doc2vec). Then, they used a Decision Tree, Random Forest, Gaussian Naive Bayes, SVM, K-Nearest Neighbors, Logistic Regression, and AdaBoost to compare six US airline businesses. 80% of the data was used to train the classifiers, and testing was conducted using leftover data. They divided the sentiments expressed in the tweets into three groups. They reported that SVM, Random Forest, AdaBoost, and Logistic Regression all had accuracy levels above 80% and performed well. According to their findings, the AdaBoost method, however, is a more reliable classifier than the others.

**LITERATURE SURVEY:**

The authors of "A Text Mining Application of Emotion Classifications of Twitter's Users Using Naive Bayes Method," Liza Wikarsa and Sherly Novianti Thahir, developed a classification model to sort the text in tweets according on sentiment polarities. The test studies shown that specific phrases and a larger training dataset obtained a greater accuracy for the identification of emotions because they may provide a better and wider coverage of the emotional situations in our daily lives. In their article "Opinion Mining of Twitter Data for Recommending Airlines Services," Pranika Jindala, Varun Jaiswala, and M. Umac examined various classification models with measures of sentiment analysis and found that the new ensemble ada boost approach had the highest accuracy value. They desire to use these models in several languages and also needs client data to add or modify the current functionality. The authors of "Tweet sentiment analysis with classifier ensembles," Nadia F.F. da Silva, Eduardo R. Hruschka, and Estevam R. Hruschka, employed ensemble classification methodologies for several classification models and compared the accuracy of the ensemble classification models. They solely exploited the polarities of positive and negative sentiment. They will apply the classification models to different datasets that are neutral in terms of sentiment.

The area of text classiﬁcation possesses a huge potential to analyse sentiments and many researchers have investigated the process of sentiment analysis by detecting emotions found in the text. Others have proposed sentiment evaluation methods that are formulated by observing human responses to a certain experience. The use of machine learning techniques including naïve Bayes (NB), maximum entropy (ME), and support vector machines (SVM) for sentiment classiﬁcation has also been studied. For example, the authors

applied NB, ME, and SVM on the Internet Movie Database (IMDb), which

consists of movie reviews expressed either with stars or in numerical values. The approach is evaluated using accuracy and recall measures. This work has served as a baseline for many authors and the same techniques have been utilized across different domains.

Similarly, the authors of performed sentiment analysis on travelers’ feedback about airlines.

The authors [1] found that the feature selection and over-sampling techniques are equally important to achieve reﬁned results. Feature analysis is performed to select the best features which not only improves the overall performance of the model but reduces the training time as well. In addition, the skewed distribution of the classes found in most of the smaller datasets is reduced without causing over-ﬁtting. The results of the research show the compelling evidence that the proposed model has a higher classiﬁcation accuracy

when predicting the three classes of positive, negative, and neutral. The authors of followed a similar approach and performed a multi-class sentiment classiﬁcation. A feature selection process is used to extract the important features that are later used to train a machine learning-based algorithm.

The authors [2] used customers feedback to investigate different aspects such

as loyalty, satisfaction, etc. The loyalty is determined through airline attributes, namely operational factors (punctuality, aircraft, and safety), attractive factors (food and beverages and the staff service), competitive factors (schedule, ticket prices, reputation, and ﬂyer program), etc.

The research concludes that the customer’s higher satisfaction can be achieved through company reputation, staff service, frequent ﬂyer program, aircraft, and punctuality. Kumar and Sebastian presented a novel approach for the

sentiment analysis of Twitter data. To uncover the sentiment, the authors extracted the opinion words combination of the adjectives along with the verbs and adverbs) in the tweets. The corpus-based method is used to ﬁnd the semantic orientation of adjectives and the dictionary-based method to ﬁnd the semantic orientation of verbs and adverbs. The overall tweet sentiment is then calculated using a linear equation that also incorporates emotion intensiﬁers. A score is calculated for the overall sentiment of the tweet and

tweets are classiﬁed as positive, neutral and negative based on the calculated score.

The authors [3] of performed sentiment analysis using a machine learning technique. The polarity is found that contains opinions extracted from a wordnet database.

presented a meta-heuristic method called CSK, which is based on cuckoo search

(CS) and k-means (K). Since clustering plays a vital role in analyzing the viewpoints and sentiments in user tweets, the research proposes a method that is used to ﬁnd the optimum cluster head from the twitter dataset. Experimental results show promising outcomes.

The authors [4] of investigated the impact of multiple classiﬁer systems on Turkish sentiment classiﬁcation. The voting algorithm is used with NB,

SVM, and bagging to evaluate their efﬁcacy. The results demonstrate that the use of multiple classiﬁers elevates the performance of individual classiﬁers. The research approves that multiple classiﬁer systems have more potential for sentiment classiﬁcation.

In addition to the use of multiple classiﬁers for classiﬁcation, employing various pre-processing techniques helps to improve the classiﬁcation as well.

The authors of [5] proved that the selection of an appropriate pre-processing technique may produce enhanced classiﬁcation performance. They investigated a variety of pre-processing techniques including term weighting, frequency cut, stemming, and stopword elimination to analyse their impact on machine learning-based classiﬁcation methods. Their research shows that the combination of various pre-processing methods plays a decisive role in ﬁnding the best classiﬁcation rates. They also studied the pre-processing techniques and their relevant impact on the feature space through visualization. In the same fashion, the use of various feature extraction techniques has proven to improve

classiﬁcation accuracy. Text mining has many feature extraction methods but term frequency.

The authors [6] investigated the use of TF, IDF, and TF-IDF with linear

classiﬁers including SVM, LR, and perceptron with a native language identiﬁcation system. Experiments are carried out with ten-fold cross-validation on different languages. The TF-IDF is applied to n-gram words/characters/ parts-of-speech tags. The TF-IDF weighting on features proves to outperform other techniques when applied with uni-grams and bi-grams.

**METHODOLOGIES:**

Data Data

Preprocessing

Stopwords removalStopwords removal

StemmingStemming

Convert to lower caseConvert to lower case

Punctuation removalPunctuation removal

Numerical removalNumerical removal

Data split

Corpus Corpus

Train Train

Test Test

Feature

engineering

TF-IDFTF-IDF

TFTF

Classifiers/Learning models

SVCSVC

SGDCSGDC

ADBADB

RFRFDTCDTC

LRLRGNBGNB

GBMGBMETCETC

Prediction

Test dataTest data

Trained modelTrained model

Evaluation

parameters

Accuracy Accuracy

Precision Precision

Recall Recall

F1-scoreF1-scor

Data Data

Preprocessing

Stopwords removalStopwords removal

StemmingStemming

Convert to lower caseConvert to lower case

Punctuation removalPunctuation removal

Numerical removalNumerical removal

Data split

Corpus Corpus

Train Train

Test Test

Feature

engineering

TF-IDFTF-IDF

TFTF

Classifiers/Learning models

SVCSVC

SGDCSGDC

ADBADB

RFRFDTCDTC

LRLRGNBGNB

GBMGBMETCETC

Prediction

Test dataTest data

Trained modelTrained model

Evaluation

parameters

Accuracy Accuracy

Precision Precision

Recall Recall

F1-scoreF1-scor

Train

Cleaning of data

Data

Collection

TF-IDF

Test

EDA

Data processing

Data Visualization

Classifiers/Models Parameters

**Figure 1: METHODOLOGY**

Predictions

Accuracy

Random Forest

SVM

Precision

Test data

Recall

Trained Model

XG-Boost

Decision tree

F1-score

Naïve bayes

**Step1:** The dataset used in this study, dubbed "Twitter US Airline Sentiment," has 15 columns and a total of 14,640 tweets. Six of the main US Airlines were represented by the tweets gathered from Twitter in February 2015: United, US Airways, Southwest, Delta, and Virgin America. The tweets included a mix of positive, negative, and neutral feelings, and they each included a confidence rating for the assigned label as well as the justification for a negative classification. The features that are included are tweet id, sentiment, confidence score in sentiment, negative reason, confidence in negative reason, airline, confidence in sentiment gold, name, retweet count, tweet text, tweet coordinates, time, date, and location of the tweet as well as the time zone of the user who posted it.

**Step2:** Punctuation was taken out of the data because it wasn't necessary for the study's text analysis. Although punctuation makes sentences more readable, it makes it harder for models to distinguish between punctuation and other letters [30]. The tweets' numerical values were eliminated in the following stage because they had no bearing on text analysis. deleting numbers from values decreases the complexity of training the models.

**Before:** @VirginAmerica I didn’t today... Must mean i need to take another trip for 2 months!

**After:** I did not today Must mean i need to take another trip for months

**Step3:** The corpus was split into a "training subset" and a "testing subset" following the pre-processing stage. For training and testing, it was split in a 3:1 ratio, respectively. The training subset was then subjected to feature extraction techniques, as depicted in Figure 6, which illustrates the employed methodology. Both the training data used to develop the selected models and the testing data used for classification both underwent feature extraction techniques.

**Step4:** TF-IDF: A common scoring method in information retrieval (IR) and summarization is TF-IDF. The goal of TF-IDF is to demonstrate how pertinent a phrase is in a particular document. TF-IDF feature extraction takes both TF and IDF into account. IDF awards tokens that are uncommon across the board in a dataset. (term frequency (TF), inverse document frequency (IDF)).

**Step 5:** Next, in this section compared bagging classifiers and non-bagging classification techniques.

The classification techniques are: Random Forest Classification, Naive Bayes , Support Vector Machine, Decision Trees , Extreme Gradient Boosting (XGB)

**Step 6:** Evaluation of parameters:

Accuracy: The percent of true categorized measurements to all actual measurements

Precision: Precision is the percentage of the true positive divided by sum of true positive and false positive.

Recall: Recall is the percentage of true text measures from the input values that were actually measured by the structure.

F1-score: f1-score measures from a weighted mean of precision and recall values.

**DATA DESCRIPTION:**

The dataset used in this study, dubbed "Twitter US Airline Sentiment," has 15 columns and a total of 14,640 tweets. Six of the main US Airlines were represented by the tweets gathered from Twitter in February 2015: United, US Airways, Southwest, Delta, and Virgin America. The tweets included a mix of positive, negative, and neutral feelings, and they each included a confidence rating for the assigned label as well as the justification for a negative classification. The features that are included are tweet id, sentiment, confidence score in sentiment, negative reason, confidence in negative reason, airline, confidence in sentiment gold, name, retweet count, tweet text, tweet coordinates, time, date, and location of the tweet as well as the time zone of the user who posted it.

| Attribute’s Name | Count | Data type | Null Values |
| --- | --- | --- | --- |
| tweet\_id | 14640 | float | 0 |
| airline\_sentiment | 14640 | category | 0 |
| airline\_sentiment\_confidence | 14640 | float | 0 |
| negativereason | 9178 | string | 5462 |
| negativereason\_confidence | 10522 | float64 | 4118 |
| airline | 14640 | category | 0 |
| airline\_sentiment\_gold | 40 | object | 14608 |
| Name | 14640 | category | 0 |
| negativereason\_gold | 32 | object | 14608 |
| retweet\_count | 14640 | int | 0 |
| text | 14640 | string | 0 |
| tweet\_coord | 1019 | object | 13621 |
| tweet\_created | 14640 | object | 0 |
| tweet\_location | 9907 | object | 4733 |
| user\_timezone | 9820 | object | 4820 |

**Table 1: Data Information**

**Analyzing Exploratory Data (EDA):**

SUMMARY OF THE NUMERICAL ATTRIBUTES:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | tweet\_id | airline\_sentiment\_confidence | negativereason\_confidence | retweet\_count |
| count | 1.46E+04 | 14640 | 10522 | 14640 |
| mean | 5.69E+17 | 0.9 | 0.64 | 0.08 |
| std | 7.79E+14 | 0.16 | 0.33 | 0.75 |
| min | 5.68E+17 | 0.34 | 0 | 0 |
| 25% | 5.69E+17 | 0.69 | 0.36 | 0 |
| 50% | 5.69E+17 | 1 | 0.67 | 0 |
| 75% | 5.70E+17 | 1 | 1 | 0 |
| max | 5.70E+17 | 1 | 1 | 44 |

**Table 2: Descriptive analytics (summary of the numerical attributes).**

Post that, I have changed the data type of above attributes according to the project need.

**Changing data type:**

For example: (airline\_sentiment ) to category

Which will further contain neutral, negative and positive. Among them most of the tweets are negative

**Replacement of null values**: Then I have replaced the null values as

For eg: "negativereason": "Unknown".

After replacement of null values. I have removed some of the attributes which I further do not want in my project.

**Data Duplicacy:**

Then I have checked the duplicated values. Post that I have removed all the duplicated values and checked the description od the data using describe().

**• DATA PREPROCESSING:** Our data have been examined for the missing value. Three columns with more than 90% empty values were removed: "tweet coord," "airline sentiment," and "negativereason." Additionally, the useless "tweet id" column, which did not advance our objective, was eliminated.

The dataset used for sentiment analysis is described in this part, along with how it was visualised and the methodology that was suggested for applying sentiment analysis to the chosen dataset.

The dataset underwent pre-processing as part of the research's methodology procedures. This stage made use of a variety of programmes and libraries, including the Natural Language Toolkit. At the pre-processing level, this study took into consideration two strategies:

**Complete pre-processing:** Data cleaning was done during thorough pre-processing to increase the learning effectiveness of machine learning models. If the data are pre-processed, machine learning models exhibit increased classification accuracy. The Python's natural language toolkit was used for the pre-processing. Punctuation, stop-words, and the use of both lower- and capital letters in tweets can all have an impact on a model's ability to learn.

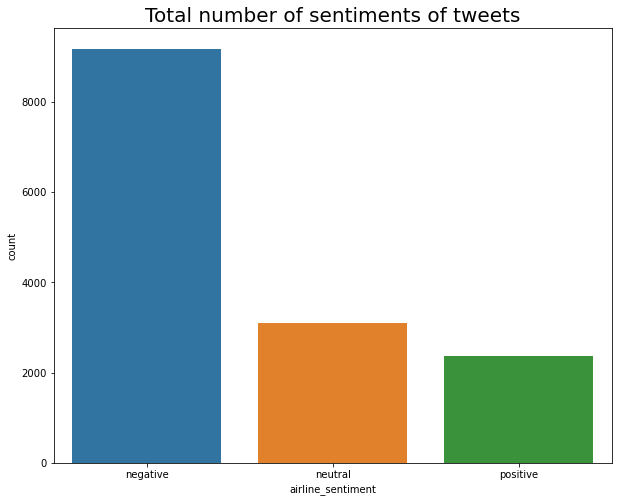
Punctuation was taken out of the data because it wasn't necessary for the study's text analysis. Although punctuation makes sentences more readable, it makes it harder for models to distinguish between punctuation and other letters. The tweets' numerical values were eliminated in the following stage because they had no bearing on text analysis. deleting numbers from values decreases the complexity of training the models.

**Before:** @VirginAmerica I didn’t today... Must mean i need to take another trip for 2 months!

**After:** I did not today Must mean i need to take another trip for months

**Partial Pre-processing:** To examine the effects of pre-processing steps on classifier accuracy, this study also took into account the usage of partial pre-processing in addition to complete pre-processing. The partial pre-processing excludes "stop-words elimination" and stemming.

**Data Visualization**



**Figure 2: Total Number of Sentiments**

Total number of sentiments of tweets

negative 9178

neutral 3099

positive 2363

This graph depicts the total number of tweets in a whole dataset and also describing the count of tweets in accordance with negative, neutral and positive.

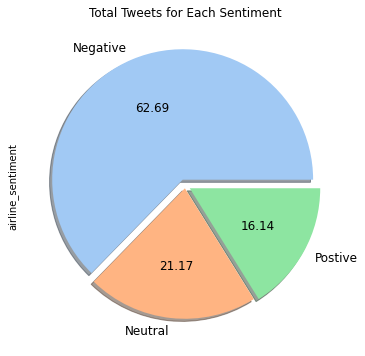
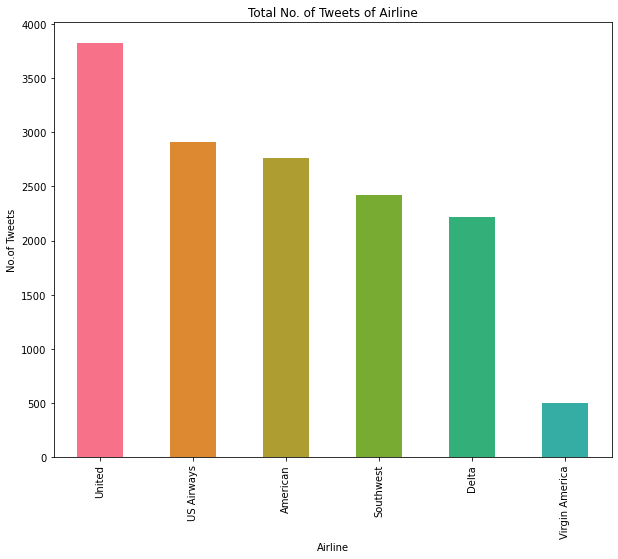


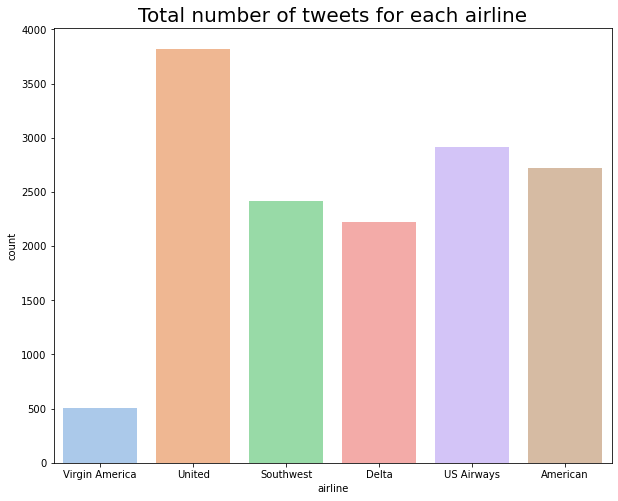
Figure 3: Total Number of Sentiments of Tweets with Percentage

The given bar-graph delineates the info about sentiments of the tweet and the percentage of tweets as from this graph, we can make conclusion that most of tweets are negative with 62.69%.



**Figure 4: Total number of negative tweets for each airline**

The bar-graph tells about the negative tweets of each airline as it depicts that highest number of tweets are for United airlines and least is for Virgin America.



**Figure 5: Total Number of tweets for each airline**

Total number of sentiment tweets for each airline :

US Airways :

negative 2263

neutral 381

positive 269

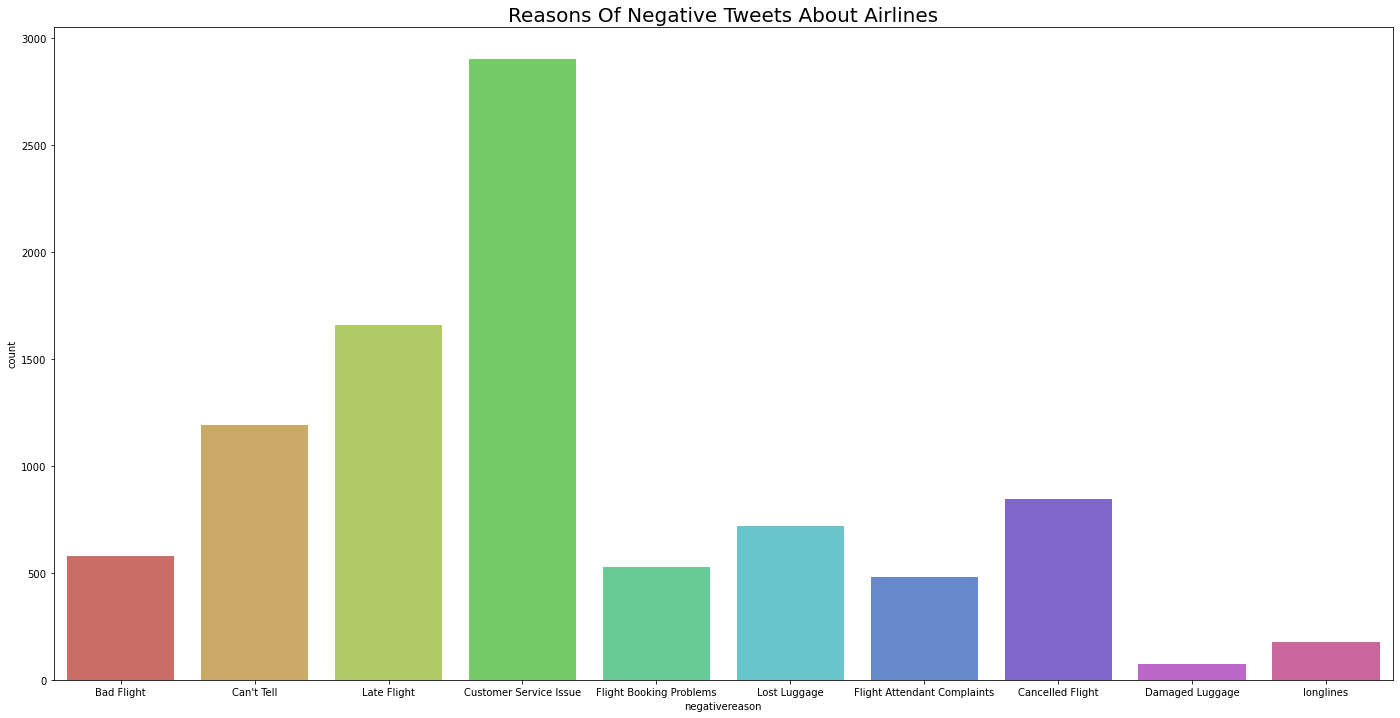
United :

negative 2633

neutral 697

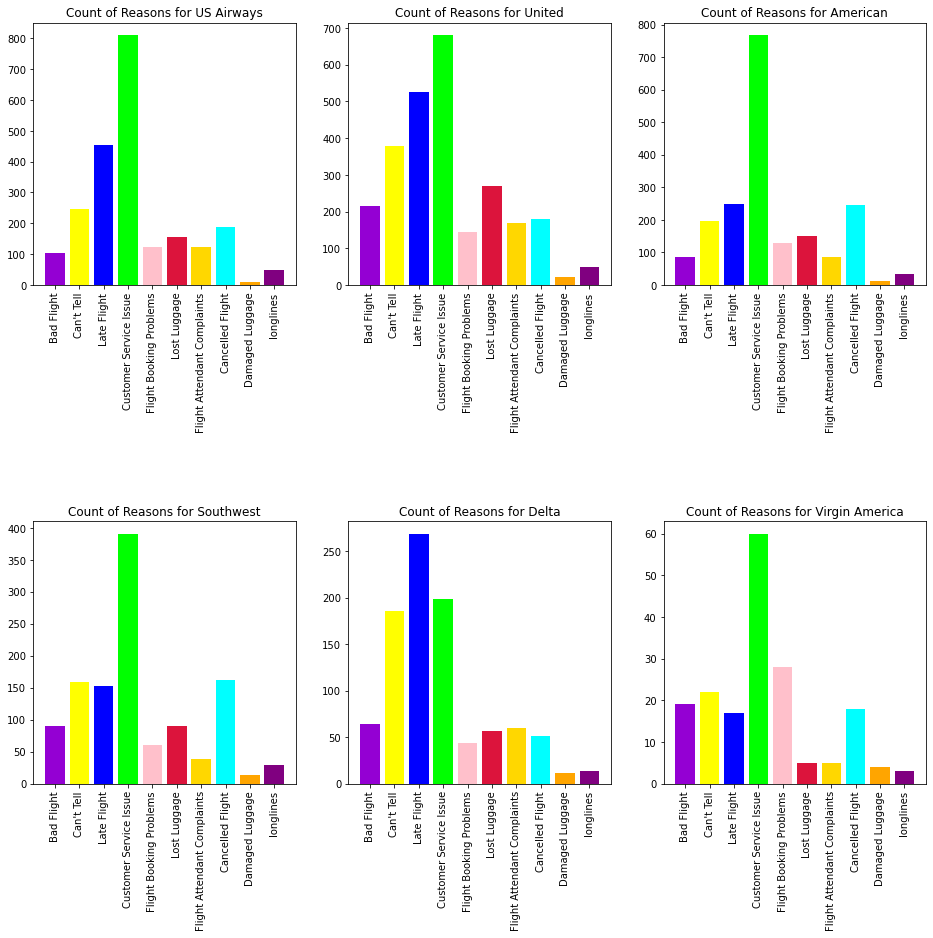
positive 492

This figure, tells about the negative tweets of each airline as well as count of positive, negative and neutral tweets. For example, US Airways has 2263-negative, 381-neutral and 269-positive.



**Figure 6: Reasons of Negative Tweets about airlines**

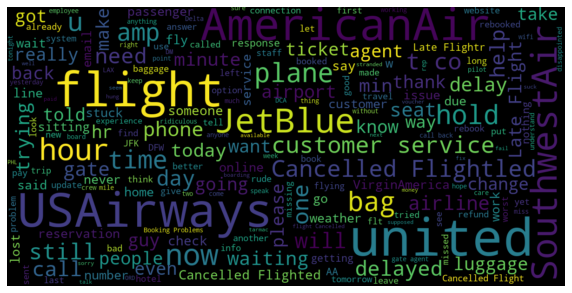
Post that, this graph tells us about the reasons of negative tweets either it is customer service issue, bad flight, late flight etcetera. Most of tweets are about customer service issue.



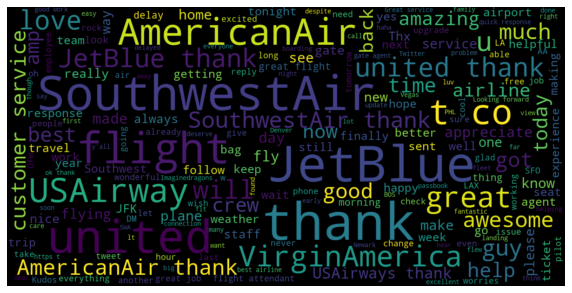
**Figure 7: Reasons of Negative tweets of each airline.**

From this figure, we can conclude that most of negative reason for each flight is about customer service issue and the least is about damaged luggage. This figure is delivering info about each airline as Us Airways, United airlines etcetera.

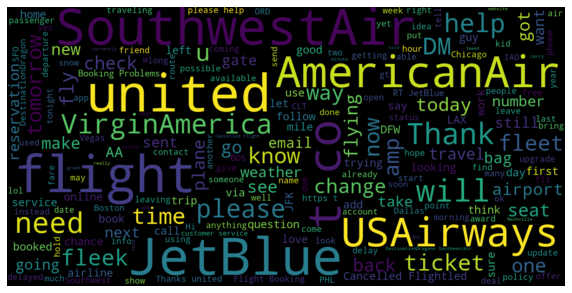
**Word Clouds:**



**Figure 8: Word cloud about the negative tweets.**



**Figure 9: Word cloud about the positive tweets.**



**Figure 10: Word cloud about the neutral tweets.**

**Feature Extraction Methods:**

**TF-IDF:** A common scoring method in information retrieval (IR) and summarization is TF-IDF. The goal of TF-IDF is to demonstrate how pertinent a phrase is in a particular document. TF-IDF feature extraction takes both TF and IDF into account. IDF awards tokens that are uncommon across the board in a dataset. (term frequency (TF), inverse document frequency (IDF)).

**Train and Test Split:** The corpus was split into a "training subset" and a "testing subset" following the pre-processing stage. For training and testing, it was split in a 3:1 ratio, respectively. The training subset was then subjected to feature extraction techniques, as depicted in Figure 6, which illustrates the employed methodology. Both the training data used to develop the selected models and the testing data used for classification both underwent feature extraction techniques.

Next, in this section compared bagging classifiers and non-bagging classification techniques.

**The classification techniques are**

* Random Forest Classification
* Naive Bayes
* Support Vector Machine
* Decision Trees
* Extreme Gradient Boosting (XGB)

**Random Forest Classification**

It is a supervised machine learning classifier because both the targets and the features are intended to forecast values. Random Forest Classification This classifier, a meta-estimator, fits a number of decision trees to various dataset samples. The model classifier's predictive accuracy is developed using the average, and over-fitting is prevented.

**Naive Bayes**

Naive Bayes classification is based on the Bayes theorem and takes the confidence between each pair of features into consideration. A limited amount of training data is required for Nave Bayes to measure the required parameters. Comparing this approach to more complex classifications, it is quick.

**Support Vector Machine**

This machine learning classification algorithm is under supervision. It is a visual representation of the training data points that have been divided into groups. The kernel approach is supported by SVM, and kernel SVM permits appliance non-linearity.

**Decision Tree**

Using a decision tree approach, one can build complex trees and make changes to the data to produce radically different results.

**Gradient Boosting**

Collaborative Classification Approach for Airline Tweets Using Extreme Gradient Boosting (XGB.A gradient-boosted decision tree operation called XG-Boost is set up for speed, accuracy, and performance. Concerns with classification and regression predictive modelling are managed by XG-Boost.

**Evaluation of parameters:**

• Accuracy: The percent of true categorized measurements to all actual measurements

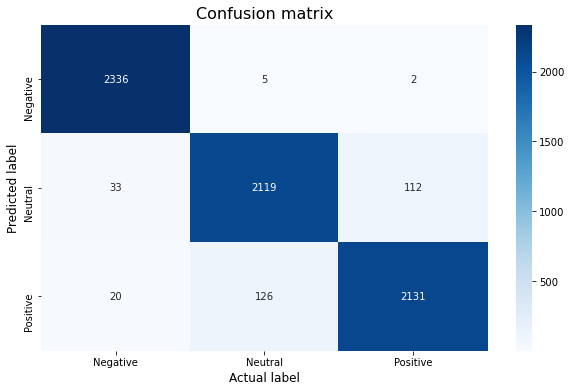
• Precision: Precision is the percentage of the true positive divided by sum of true positive and false positive.

• Recall: Recall is the percentage of true text measures from the input values that were actually measured by the structure.

• F1-score: f1-score measures from a weighted mean of precision and recall values.

| Models | Accuracy | Precision | F1- score |
| --- | --- | --- | --- |
| Random Forest  Classification | 0.9529 | 0.98 | 0.99 |
| Naïve Bayes | 0.8593 | 0.90 | 0.92 |
| SVM | 0.83651 | 0.92 | 0.91 |
| Decision Tree  Classifier | 0.9417 | 0.93 | 0.96 |
| XG-Boost | 0.8720 | 0.89 | 0.90 |

**Table 3: Table of Parameters.**

****

**Figure 11: Confusion Matrix for random forest classifier.**

|  | precision | Recall | F1-score | Support |
| --- | --- | --- | --- | --- |
| Negative | 0.98 | 1.00 | 0.99 | 2343 |
| Neutral | 0.94 | 0.94 | 0.94 | 2264 |
| Positive | 0.95 | 0.94 | 0.94 | 2277 |
| Accuracy |  |  | 0.96 | 6884 |
| Macro average | 0.96 | 0.96 | 0.96 | 6884 |
| Weighted average | 0.96 | 0.96 | 0.96 | 6884 |

**Table 4: Classification Report of Random Forest Classifier.**

**CONCLUSION:**

To conclude the whole survey or project, as my dataset is “US Airline Twitter Sentiments”, as the name described that its about the sentiments of the people regarding their experience while travelling in the flights. The dataset contains 14640 records and 15 attributes from which have dropped some attributes and changed the datatype according to the need as well as replaced the null values. Post this, by removing stop-words and punctuations from the ‘text’ attribute, all in all text pre-processing or text-analysis is done based on the requirement. Then, to normalize the data TF-IDF vectorizer is used and for handling imbalance SMOTE is used.

Post that, data visualization is also done based on the research papers studied. Word cloud about the negative, neutral and positive is described as well.

In terms of research questions, the findings are:

* Most of the flights got negative tweets about the customer service issue. So, they have to focus on this issue while, the least filed issue is damaged luggage.
* Most tweeted complaints are about Customer Service issue.
* Mostly flights had negative tweets which I had classify all the issues using visualization.

So, last but not least few methodologies are used as random forest classifier, decision tree, SVM etcetera to find the best model to predict the future. In accordance with accuracy the best fit model is “Random Forest” with 96% accuracy.

***References:***

*1. Jacobson, R. 2.5 Quintillion Bytes of Data Created Every Day. How Does CPG & Retail Manage It; IBM: Indianapolis, IN, USA, 2013.*

*2. Wang, Q.; Kealy, A.; Zhai, S. Introduction for the Special Issue on Beyond the Hypes of Geospatial Big Data: Theories, Methods, Analytics, and Applications. Comput. Model. Eng. Sci. 2019, 119, 245. [CrossRef]*

*3. Pang, B.; Lee, L. Opinion mining and sentiment analysis. Found. Trends Inf. Retr. 2008, 2, 1–135. [CrossRef]*

*4. Rainie, L.; Horrigan, J. Election 2006 online. In Pew Internet & American Life Project Report; Pew Research Center: Washington, DC, USA, 2007.*

*5. Qaisi, L.M.; Aljarah, I. A twitter sentiment analysis for cloud providers: A case study of Azure vs. AWS. In Proceedings of the IEEE 2016 7th International Conference on Computer Science and Information Technology (CSIT), Amman, Jordan, 13–14 July 2016.*

*6.Liza Wikarsa, SherlyNoviantiThahir, “A Text Mining Application of Emotion Classifications of Twitter’s Users Using Naïve Bayes Method” .*

*7E. Prabhakar and K. Sugashini, “New Ensemble Approach to Analyze User Sentiments from Social Media Twitter Data”,*

*8The SIJ Transactions on Industrial, Financial & Business Management (IFBM), Vol. 6, No. 1, 2018. Adeborna, Esi, and KengSiau. 2014. "An approach to sentiment analysis-The case of airline quality rating." Pacific Asia Conference on Information Systmem.*

*9Nadia F.F. da Silva, Eduardo R. Hruschka, Estevam R. Hruschka, "Tweet sentiment analysis with classifier ensembles” Science direct, 2014.*

*10S. Kumar and M. Zymbler, “A machine learning approach to analyze customer satisfaction from airline tweets,” Journal of Big Data, vol. 6, no. 1, 2019.*

*11. D. D. Das, S. Sharma, S. Natani, N. Khare, and B. Singh, “Sentimental Analysis for Airline Twitter data,” in IOP Conf. Series: Materials Science and Engineering, 2017.*

*12. H. Hakh, I. Aljarah, and B. Al-Shbou, “Online Social Media-based Sentiment Analysis for US Airline companies,” 2017.*