International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/IJSRCSEIT

Airlines Twitter Sentiment Analysis Using EDA AND ML

Amey Pathare , Ajay Nishad, Saurabh Dubey, Shankar Malve, Prof. Uma Goradiya
Department of Computer Engineering, Shree LR Tiwari College of Engineering, Mumbai, Maharashtra, India

ARTICLEINFO

Article History:
Accepted: 01 April 2023
Published: 22 April 2023

Publication Issue
Volume 10, Issue 2
March-April-2023

Page Number
519-523

ABSTRACT

Today, social media is a part of everyone, so there are a lot of user views available. A sentiment analysis technique called "airlines sentiment analysis" may be used to examine the issues or views of all significant U.S. airlines. Twitter data from various years was scraped in order to understand the customer's voice. Contributors were requested to first categorise positive, neutral, and negative tweets before classifying negative tweets like delayed flights or rude service. The collection of sentiment tweets for six US airlines, whether they were positive, negative, or neutral, is included in the CSV file. Social media sites like Facebook, Twitter, etc. are being used more frequently by people worldwide. Using social media, they exchange ideas, views, and information. More attention was paid to these social issues by the business

Keywords : Pertaining, Gauge, Tokenizing, Sentiment

I. INTRODUCTION

It has been determined that the need for sentiment analysis via Twitter is a more interesting area for research because it offers simpler sentiments of the general public, particularly in the case of the airline industries where people are satisfied and dissatisfied very easily and more frequently express their views on Twitter. The amount of information present in Twitter is unfathomably large compared to other data sources like review websites or blogs, where the information discovered is not only fully informative but also devoid of numbers. Despite appearing tiny, Twitter has 100 million users, or about 1/60th of the world's population. Each day, more than half a billion tweets are sent there, and the number keeps increasing with each passing day. In order to keep their market position in the fiercely competitive airline business, airlines must ensure customer satisfaction and loyalty. Customers now frequently share their thoughts, experiences, and comments about airline services on social media sites like Twitter. In order to understand customer satisfaction and pinpoint areas for development, the airline industry has turned to sentiment analysis of tweets about airlines on Twitter as a key tool. The process of automatically identifying and classifying the views stated in a piece of text as positive, negative, or neutral is known as sentiment analysis. Sentiment analysis of tweets in the context of airlines can reveal information about customer satisfaction levels, spot recurring themes or subjects linked to positive and negative sentiment, and assist airlines.

Copyright: © the author(s), publisher and licensee Technoscience Academy. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited

II. PROBLEM STATEMENT

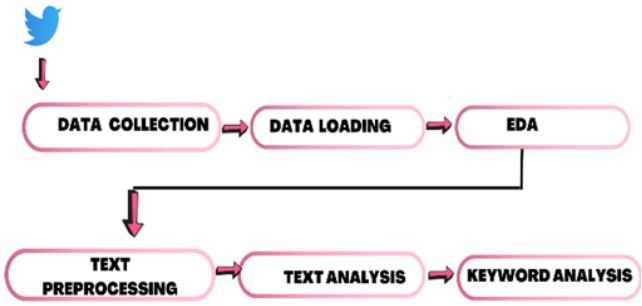
Travelers judge airlines fairly or unfairly based on a variety of experiences, some of which the airlines can influence and some of which they can't. However, if passenger has a bad trip, there's just as likely that it was a result of the lengthy line at the airport security checkpoint, the icy weather, or an outdated terminal as it was of a lost bag or an overbooking. As 99% of customers use social media sites like Twitter to share their opinions about their experiences. Whether the text is conveyed as a sentence or as an entity, either a positive, negative, or neutral aspect or feature will be present

III. OBJECTIVES

- To build an online system for sentimental analysis for airlines twitter data.
- To classify the tweets as positive, objective/neutral and negative.
- To examine the causes of each airline's negative comments.
- To better understand brands, increase impact, and enhance customer service
- To measure the level of customer satisfaction with different airline services, such as punctual arrivals, hospitable staff, in-flight comforts, etc.
- To assess the tone of tweets about different airline companies and compare it to those of competitors.

IV. METHODOLOGY

Punctuation and emojis were left in because they symbolise certain emotions in the first stage, which also involved tokenizing the data and removing words like URLs and numbers. Finding opinions required the use of the pre-processed data. After removing punctuation and emojis, sentiments were calculated using the same feature as in the first stage, but without the use of emoticons. Tokenizing the tweet data from the carriers serves as the first step in the suggested algorithm. The information is obtained from Twitter and saved as.csv files. The system will then be put through two stages of operation: the Training phase, during which it will be taught using data, and the Testing phase, during which it will be put through its paces to produce the desired outcome in real time.



WORKING OF THE SYSTEM

V. LITERATURE REVIEW

The tone of comments about airlines on Twitter has been examined in a number of studies. They used machine learning to analyse the sentiment of tweets pertaining to significant US airlines. According to the research, tweets generally had a negative tone, with flight delays and cancellations being the most frequently voiced complaints. Each tweet was categorised by the authors into one of three sentiment categories: positive, negative, or neutral using a technique called natural language processing. To gauge the strength of the sentiment expressed in each tweet, they also performed a sentiment intensity analysis. According to the survey, 21.6% of tweets were favourable, 34.9% were negative, and 43.5% were neutral. In order to determine the subjects that appeared most frequently in the tweets about the three airlines, the authors also performed a topic analysis. According to the report, baggage handling, customer service, and aircraft delays and cancellations were the most commonly discussed subjects. [1]. The sentiment of tweets about Chinese airlines was usually positive, with airline promotions and customer service being the most frequently

discussed subjects. In e-commerce, transportation, and industry, natural language processing methods are used. [2]. For instance, these colleagues used machine learning to analyse the sentiment of tweets about Indian airlines. The authors gathered a dataset of 8,863 tweets that mentioned five significant Indian airlines: Air India, IndiGo, SpiceJet, Jet Airways, and GoAir. These tweets were published between April and May 2016. The Support Vector Machine (SVM) algorithm, a well-known method in natural language processing, was used by the authors to do sentiment analysis using machine learning methods. Stop words, special characters, and URLs were removed during pre-processing of the tweets by the writers. Additionally, they used stemming to break words down to their root form. According to the survey, most tweets had a neutral sentiment, with 12.2% having negative sentiment, 30.4% having positive sentiment, and 57.4% having neutral mood. The authors also looked at the tweets about Indian airlines that were most commonly discussed..The research discovered that there was a generally negative attitude, with flight cancellations and delays being the most frequently reported complaints. [3]. For instance, they looked at the connection between customer loyalty to big US airlines and sentiment on Twitter. The research discovered a positive correlation between the sentiment of tweets and customer loyalty, indicating that airlines can use social media to increase customer loyalty. [4]. The customer reviews of Chinese airlines were analysed using sentiment analysis. The six main Chinese airlines Air China, China Eastern, China Southern, Hainan Airlines, Spring Airlines, and Xiamen Airlines were the subjects of the research methodology's collection and analysis of online customer reviews. Nine,042 reviews in all were gathered from four significant web directories, including TripAdvisor, Qunar, Ctrip, and Douban. A Chinese sentiment lexicon was used for the sentiment analysis, classifying each review as either positive, negative, or neutral. The overall sentiment score for each airline was then determined using the findings of the sentiment study. The survey, which was based on an emotion dictionary, discovered that customers were more worried about flight on-timeness, service quality, and ticket prices. [5]. Sentiment analysis was used to examine customer reviews of Saudi Arabian airlines.The study used a lexicon-based methodology and discovered that customers were mostly worried about service quality, on-time flights, and employee attitude. The TextBlob library, a well-known natural language processing programme, was used by the authors to undertake sentiment analysis. Each tweet was categorised by the study as having a favourable, negative, or neutral emotion. According to the survey, 60.5% of tweets were favourable, 15.8% were unfavourable, and 23.7% were indifferent. The authors also looked at the tweets about Saudi Arabian airlines that were most frequently discussed. According to the report, the most often stated subjects were luggage handling, in-flight services, and flight delays and cancellations. Positive mood was linked to in-flight amenities, while negative emotion was linked to flight delays and cancellations, according to the sentiment research. [6]

VI. PROPOSED SYSTEM

In this study, tweets from Twitter about people's flying experiences are collected, and airline firms can offer suggestions based on the travellers' views. About 14,535 tweets total, gathered in February 2015 from various airline evaluations, make up the dataset.About 86% of the unfavourable reviews in our data set have a marked negative reason, while the remaining 14% are classified "can't tell." Reviews are categorised as either good, negative, or neutral. First, each negative response's cause was examined, such as a delayed flight or misplaced luggage, and seven distinct classifier models were developed to do sentiment analysis on the data set. Knowing the

rationale behind every unfavourable review allows us to provide targeted recommendations.

VII. PROJECT SCOPE

- There are a lot of exciting possibilities offered by Twitter sentiment analysis. It gives social media tracking a new perspective by being able to analyse tweets in real time and ascertain the emotion behind each message.
- You can monitor sentiment using Twitter Sentiment Analysis, which offers fascinating insights into how people feel about a particular candidate.
- Numerous texts can be automatically and rapidly identified and analysed using sentiment analysis tools.
- However, computer programmes struggle to understand things like sarcasm and irony, negations,
- jokes, and exaggerations. For example, if someone wrote, "Disappointed," it might be classified as a negative word for sentiment analysis purposes, but if they meant, "Wasn't Disappointed," it would be classified as a positive statement.

VIII. RESULTS AND DISCUSSION

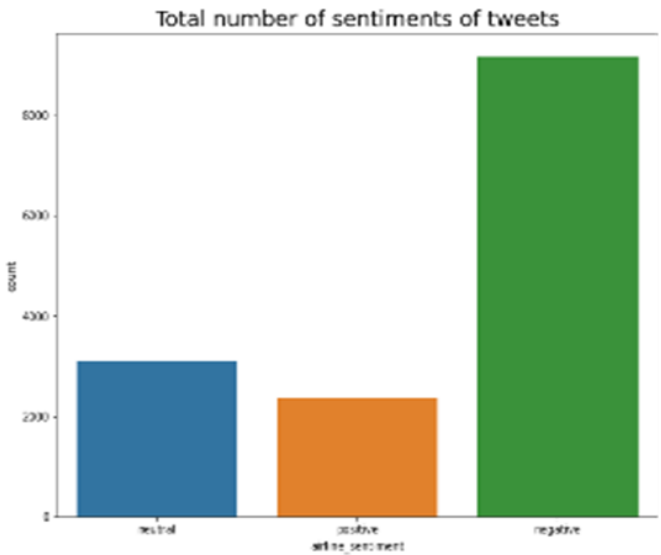


Fig 1

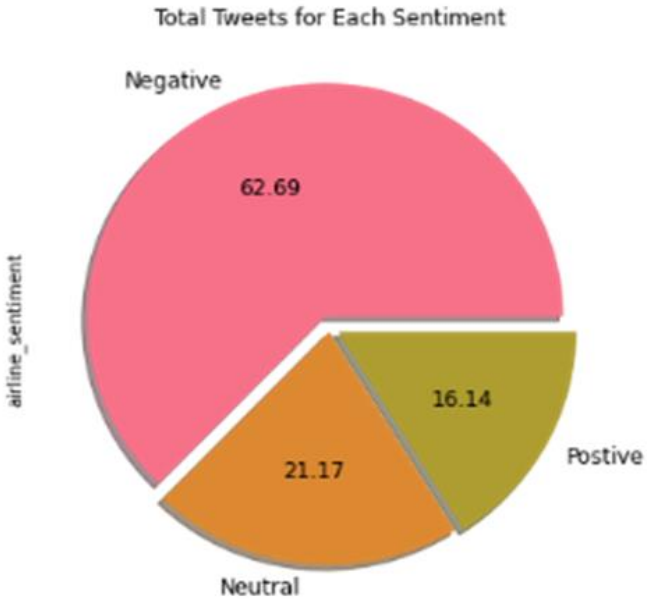


Fig 2

With the aid of random forest classification, the results from the twitter data spreadsheet showed that 1269 out of the 22375 tweets for all the airlines were classified as being negative, 21.17 as neutral, and 16.14 as positive. Seeing the Fig(1) we can say that the negative tweets were tweeted higher as compared to neutral and positive. The positive tweets has been tweeted lower.

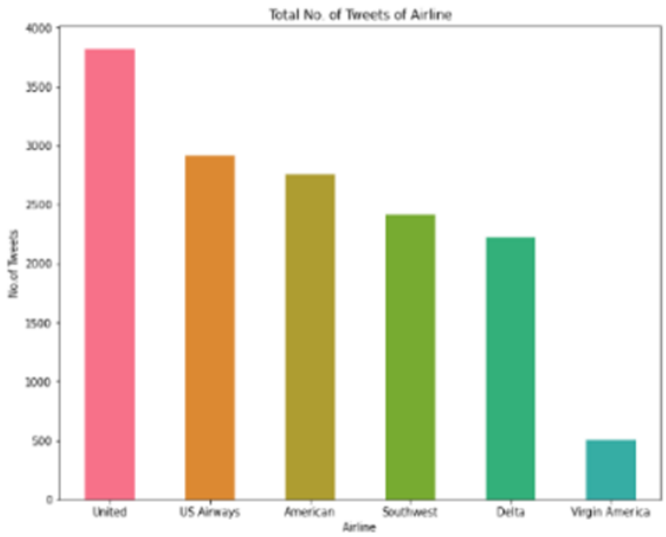


Fig 3

`The Total no. of tweets for united airlines are more compared to other. This can include positive, negative and neutral followed by the lowest no. of tweets are less for Virgin America Airlines.

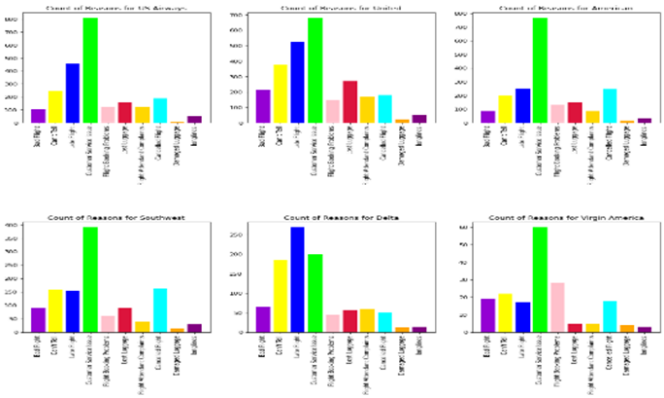


Fig 4

The most negative tweets about poor customer care are directed at US Airways, United, American Airlines, Southwest, and Virgin America. While the late departure on Delta Airlines is the cause of more tweets. As a result, US Airways, United, American Airlines, Southwest, and Virgin America are in competition with one another to provide better services, stick out from the crowd, and draw more customers.

IX. CONCLUSION AND FUTURE WORK

Nowadays, social media is crucial for marketing and presenting information like opinions and evaluations. With the development of technology, there is an enormous quantity of data on social media that is growing in volume, subjectivity, and difficulty to manually process. This sentiment analysis system for airline tweets uses a variety of machine learning techniques that will be created and implemented after reading up on the subject in the literature. Therefore, in the future, we must create a classifier that correctly and efficiently addresses customer and consumer problems. The analysis can assist airlines in creating strategies that will increase client loyalty and happiness. Sentiment analysis findings should be interpreted cautiously, though, as tweets might not accurately capture consumer sentiment as a whole. Exploring the connection between online sentiment and offline consumer satisfaction will require more study.

X. REFERENCES

[1]. X. "Mining airline passenger sentiment from Twitter data," Journal of Air Transport Management, pp. 10-17, 2017.

[2]. Wang, "Analysis of airline user-generated content on social media: Sentiment, topics, and insights," Journal of Air Transport Management, pp. 49-61, 2018.

[3]. Nair and S. , "A sentiment analysis of tweets related to Indian airlines using machine learning techniques," Journal of Air Transport Management, pp. 101-110, 2017.

[4]. Lee, "The influence of Twitter sentiment on customer loyalty in the airline industry," Journal of Air Transport Management, pp. 135-143, 2018.

[5]. Alharbi and A. , "Analysis of customer sentiment towards Saudi Arabian airlines using sentiment analysis," Journal of Airline and Airport Management, pp. 1-18, 2019.

[6]. Guo, B. and J. , "Chinese airline passenger satisfaction and loyalty: An analysis of online reviews," Journal of Air Transport Management, pp. 99-108.

Cite this article as :

Amey Pathare, Ajay Nishad, Saurabh Dubey, Shankar Malve, Prof. Uma Goradiya, "Airlines Twitter Sentiment Analysis Using EDA AND ML", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 9, Issue 2, pp.519-523, March-April-2023.