JamesMbeti /
AIRLINE-ENTIMENT-ANALYSIS

<> **Code** · Issues ⑆ Pull requests ▷ Actions ▦ Projects 📖 Wiki ⓘ Security Insights ⚙

👁 ⑃ ☆

☆ **0** stars ⑃ **0** forks 👁 **1** watching ⎵ Activity

🌐 Public repository

⑃ main ▾ ···

⑃ Branches ◇ Tags

Your main branch isn't protected

Protect this branch from force pushing or deletion, or require status checks before merging. Learn more

Protect this branch   Dismiss

JamesMbeti ···   3 hours ago ⟳

View code

☰ **README.md** ✎

AIRLINE-SENTIMENT-ANALYSIS

BUSINESS OVERVIEW

INTRODUCTION

The airline industry is highly competitive, and customer satisfaction plays a crucial role in determining the success and reputation of airlines. In today's digital age, social media platforms have become a significant avenue for customers to express their opinions and experiences with airlines. This project aims to develop a comprehensive solution for monitoring, analyzing, and understanding customer sentiment expressed on Twitter regarding various airlines. The project focuses on leveraging natural language processing (NLP) and machine learning techniques to classify tweets into positive, negative, or neutral sentiment categories. By analyzing sentiment, airlines can gain actionable insights to enhance customer satisfaction, identify operational improvements, and effectively manage their brand reputation on Twitter.

METHODOLOGY

The project will follow a structured methodology encompassing the following steps:

- Data Collection: The data was sourced from here. It was scraped from February 2015 and contains tweet reviews of different US airline companies.
- Data Preprocessing: Clean and preprocess the tweet data by removing noise, irrelevant information, and performing tasks such as tokenization, stemming, and removing stopwords.
- Sentiment Classification Model: Train a machine learning model (such as a supervised classifier or deep learning model) using the preprocessed dataset to classify tweets into positive, negative, or neutral sentiments. Evaluate the model's performance using appropriate evaluation metrics.
- Real-time Monitoring System: Implement a system that retrieves live tweets related to airlines and applies the sentiment classification model to categorize them in real time. Handle high volume and velocity of incoming tweets efficiently and ensure scalability.
- Insights and Recommendations: Analyze sentiment analysis results to generate actionable insights and recommendations for improving customer satisfaction, addressing pain points, and managing brand reputation effectively.
- Response and Engagement Strategy: Develop a strategy for airlines to respond to negative sentiment and engage with customers in a timely and personalized manner. Implement systems and processes to manage customer feedback, complaints, and turnaround negative experiences into positive ones.

PROBLEM STATEMENT

The airline industry is currently facing a notable decrease in customer satisfaction, leading to unfavorable brand perception and diminished customer loyalty. This decline in satisfaction can be attributed to several factors, including flight delays, inadequate customer service, mishandling of luggage, and other operational inefficiencies. As a result, addressing these customer concerns and enhancing the overall brand perception has become a crucial focus for airlines.

OBJECTIVES

1. Determine the overall sentiment expressed in tweets related to US airlines. This involves classifying tweets as positive, negative, or neutral to understand the general sentiment of customers towards different airlines.
2. Implement a real-time monitoring system to continuously capture and process tweets related to airlines from Twitter.
3. Generate actionable insights and recommendations based on sentiment analysis to improve customer satisfaction, address pain points, and enhance overall brand reputation.
4. Establish an effective response and engagement strategy to manage negative sentiment, address customer complaints, and foster positive customer experiences.

DATA UNDERSTANDING

The data was sourced from here. It was scraped from February 2015 and contains tweet reviews of different US airline companies.

FINAL MODEL

Our final and preferred model is the untuned random forest classifier which performed better compared to other models with an accuracy of 86.51 percent. This is a remarkable improvement from the baseline model by approximately 13%.

CONCLUSIONS

- Sentiments expressed by customers play a significant role in their decision to continue or discontinue their relationship with an airline as we've seen through the prediction of churn rates in our dataset.
- From the dataset, we are able to predict that 7% of the customers are likely to churn.
- From our sentiment analysis in the tweets, we find that most customer pain points are about canceled, delayed and missed flights.
- We settled on the final model, which is the Random Forest Model which had the highest accuracy score as compared to other models.
- The model has proven to have an accuracy of 86% in classying whether a tweet is postive, negative or neutral and it can be used to continuosly monitor the sentiments coming from the social media platforms.

MODEL LIMITATIONS

- The model incorrectly identifies one out of seven reviews.
- We had a class imbalanced and so the model maybe biased towards the majority class.

RECOMMENDATIONS

- Airlines with a higher count of negative sentiments should pay attention to the feedback provided by customers. Negative sentiments can indicate areas where improvements are needed, such as customer service, flight punctuality and baggage handling.
- Airlines should Provide comprehensive training to airline staff, including customer-facing employees such as flight attendants, ground staff, and customer service representatives in order to ensure customers are proper handled to curb on the churning rates.
- Following many complaints coming from the tweet sentiments revolving around cancaled, delayed and missed flights, the airlines should provide effective schedules and effeciency in operations on their flight depatures and incase of any challenge like bad weather, there should be proper communication to the customers in due time to avoid inconveniences.
- Personalized Marketing and Offers could help mitigate negative reviews.
- Identifying influential individuals or social media accounts within the customers through collaborative promotional campaigns.

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

Languages

● **Jupyter Notebook** 100.0%