

Proposal for Sentiment Analysis of US Airline on Twitter

To: Head of Analytics, United Airlines

From: Fred Chen, Shiyu Hua, Yifei Fan, Heidi Lu, Yiting Zhang,
Associate, United Airlines

1. Introduction

According to a new J.D. Power survey in 2019, which was based on responses from nearly 6,000 passengers who flew on major North American airlines between March 2018 and March 2019, United Airlines remained the worst-ranked traditional airline when it comes to consumer satisfaction (Zumbach, 2019). As a data scientist team in United Airline, we plan to fetch data from public Twitter pages, and extract the airline-related posts to further explore customers' reviews by using sentiment analysis. Dividing the related tweets into positive, negative and neutral ones, we can understand what are the most common reasons in case of a bad flight with UA, and how our competitors perform. Therefore, we could improve our services based on the insights of this project.

2. Challenge/Opportunity

With the recent announcement of ordering 270 new aircrafts, United Airlines expects a complete rebound in air travel (Aratani, 2021). The entire U.S. Airlines are starting to position themselves for a travel revival. However, the notable scandals and bad publicity about United Airlines has been a significant obstacle. Last month, we even faced pressure from the Department of Transportation for keeping thousands of passengers stuck on planes for hours (Duncan, 2021). To avoid losing our customers to competitors, it is urgent to put some effort now into improving our passenger experience.

Therefore, we propose to analyze the reviews of U.S. airlines posted on social media to find out the cause of dissatisfaction from our current or potential customers. Through sentiment analysis, we can identify what our customers value the most in our service. At the same time, we can also expect to better understand our strengths and weaknesses compared to our primary competitors. Ultimately, we want to improve our customer satisfaction rate and expand our loyalty base by leveraging the results of analytics.

3. Rationale

Using twitter provided API, we could easily collect United Airline and its competitors' customer tweets. Once the tweets are downloaded, they will be pre-processed and then be converted into a document-term matrix for sentimental analysis. First question we want to know is our customers' attitudes towards our airlines.

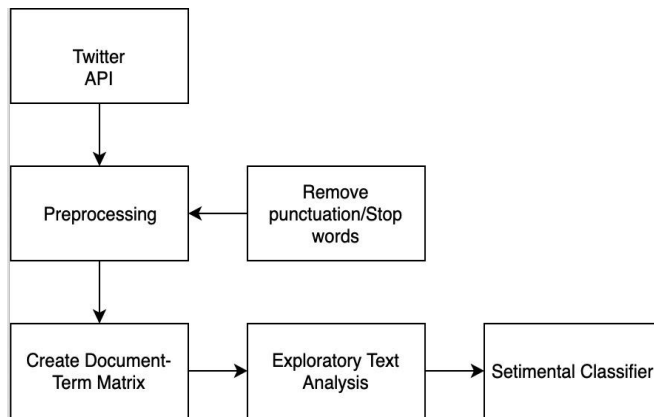


Figure 1. Sentiment analysis procedure

A sentiment classifier (using Naive Bayes) can classify tweets into three categories: positive, negative and neutral. This is the very first step to know how our customers are feeling towards our airlines. Here are some preliminary experiment results we did.

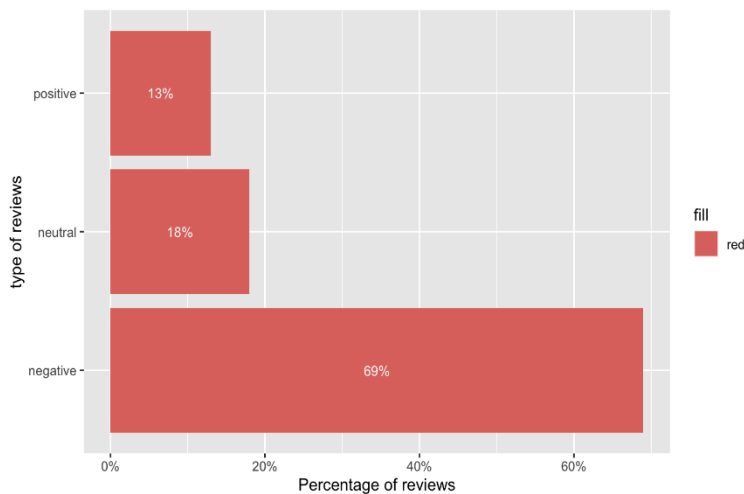


Figure 2. 69% of tweets are negative

Another question we want to solve is what are the most common reasons in case of a negative review with United Airlines. We plan to use the Latent Dirichlet allocation topic model to extract negative reasons from reviews. We've performed some experiments to draw insights from our data.

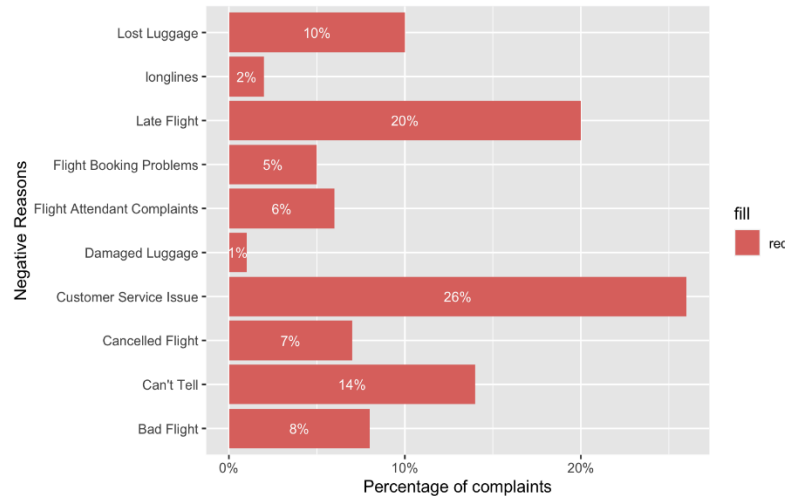


Figure 3. Top 3 complaints came from Customer Service Issue, Late Flight and unknown reasons

4. Cost Analysis

We estimate the project to last for about 1-2 months. The primary tasks include data training, language processing, and sentiment analysis. The team will consist of 3 data scientists who have expertise in natural language processing (NLP), machine learning, and hard-core coding skills (e.g., Python, SAS). Since the company is partnered with Twitter, we will not spend extra dollars on gathering the data. We estimate the total cost for this project to be around \$30,000. For value, we should compare cost to the expected gain of improving customer satisfaction and reducing customer churn rate. Customer insights provide a clear picture of our service problems and weaknesses. We can use the insights to deliver better customer experience and keep brand appreciation on track.

5. Assessment

We acknowledge that price is always the major competitive factor for airline tickets. However, our insights can assist the company to become more competitive among the established legacy carriers, Delta and American Airlines, except price. This project provides an insight on improving customer service for United Airlines relying on the Twitter US Airline Sentiment data. We can identify the part of flight experience, pre-flight, in-flight, or post-flight, that creates the most negative sentiments. In the long run, the Sentiment data becomes a reference for reflecting the group of customers who are willing to share feedback on social media. For example, identifying customer segmentation and implementing services that avoid negative sentiments in the past.

6. Conclusion

By detecting the sentiments in the tweets, our team would further explore the factors that may help us improve the customer satisfaction. Next step, we will visualize our research progress on dashboards.

7. Reference

- Aratani, L. (2021, June 29). *United Airlines bets on a pandemic recovery with order that includes 200 Boeing Max Jets*. The Washington Post. Retrieved from <https://www.washingtonpost.com/transportation/2021/06/29/united-airlines-airplane-order/>.
- Duncan, I. (2021, September 24). *United Airlines fined for keeping thousands of passengers stuck on planes for hours*. The Washington Post. Retrieved from <https://www.washingtonpost.com/transportation/2021/09/24/united-tarmac-delay-faa-fine/>.
- Zumbach, L. (2019, May 29). *United last on J.D. Power Customer Satisfaction Survey, but the airline is improving*. [chicagotribune.com](https://www.chicagotribune.com/business/ct-biz-united-airlines-ranked-last-jd-power-20190529-story.html). Retrieved from <https://www.chicagotribune.com/business/ct-biz-united-airlines-ranked-last-jd-power-20190529-story.html>.