Airline Customer Experience Analysis

Project – 2: Team Aviators

Team Members:

1. Hussain Kaide Johar Manasi (hmanasi1@student.gsu.edu)
2. Vaishnavi Mada (vmada1@student.gsu.edu)
3. Srujani Mareddy (smareddy1@student.gsu.edu)
4. Surya Kiran Varma Vegesna (svegesna1@student.gsu.edu)
5. Jinendra Shrimal (jshrimal1@student.gsu.edu)

Manager: Hussain Kaide Johar Manasi (hmanasi1@student.gsu.edu)

Statement of Academic Honesty:

The following code represents our own work. We have neither received nor given inappropriate assistance. We have not copied or modified code from any source other than the course webpage or the course textbook. We recognize that any unauthorized assistance or plagiarism will be handled in accordance with Georgia State University's Academic Honesty Policy and the policies of this course. We recognize that our work is based on an assignment created by the Institute for Insight at Georgia State University Any publishing or posting of source code for this project is strictly prohibited unless you have written consent from the Institute for Insight at Georgia State University.

Link to the google colab with complete outputs:

https://colab.research.google.com/drive/14HTJDw3GiIObVC4tXnkVuf12HZ-3Fnjd?usp=sharing

1. Introduction:

The airline industry is a pivotal part of global transportation, shaping travel experiences for millions of passengers daily. The quality of service offered by airlines significantly influences customer satisfaction, loyalty, and the industry's overall competitiveness. Understanding and improving the customer experience has become paramount in this highly competitive market.

Problem Statement:

In the realm of airline operations, gauging and enhancing customer experience is a multifaceted challenge. This analysis centers on leveraging available data – a combination of written reviews and numerical ratings – to discern patterns, sentiments, and critical factors that contribute to customer satisfaction or dissatisfaction. This report aims to delineate the intrinsic complexities of assessing airline customer experiences, particularly focusing on aspects like food and beverage services, wifi connectivity, cabin services, among others. Additionally, an ML model should allow us to build a recommendation system to any user.

Importance and Relevance:

In today's digital age, customers express their opinions and feedback on various platforms, which often include detailed reviews and quantitative ratings. This feedback can be in the form of surveys or in the form of outspoken opinions on social media. These evaluations encompass a spectrum of subjective experiences, from in-flight amenities to service quality, influencing an airline's reputation and consumer choices. The amalgamation of textual and numerical data offers a unique opportunity to unravel insights that can guide airlines in refining their services and retaining customer loyalty. If an airline can successfully cement themselves positively with a large audience, they can rely on returning customers regardless of any minor fumbles or organizational changes.

Significance of the Problem:

The significance of this problem lies in its impact on business performance. Satisfied customers tend to be loyal and can positively influence others, acting as brand advocates. Conversely, negative experiences shared widely can detrimentally affect an airline's reputation and bottom line. Therefore, decoding the nuances within customer feedback becomes imperative for airlines to tailor their services effectively, thereby potentially attracting more patrons and elevating their competitiveness in the market.

Monetary value of this Problem Statement:

Case Study 1: United Airlines (2017 Incident)

United Airlines faced a significant reputational crisis in 2017 due to an incident that went viral on social media. A passenger was forcibly removed from an overbooked flight, leading to widespread public outrage and negative publicity. The incident sparked intense backlash against United Airlines, with the public criticizing its handling of the situation, leading to a considerable drop in the airline's reputation.

Specifically, the mishandling of the situation, the lack of empathy towards the passenger, and the delayed response in addressing the issue contributed to severe damage to United Airlines' reputation. This incident highlighted the importance of customer service, crisis management, and the power of social media in shaping public perception.

Case Study 2: Singapore Airlines

Singapore Airlines stands as an industry leader renowned for its exceptional service and commitment to passenger satisfaction. It has consistently ranked among the world's best airlines due to several key factors:

1. **Service Excellence:** Singapore Airlines is acclaimed for its unparalleled customer service, emphasizing personalized and attentive care that sets industry standards.
2. **Cabin Comfort:** Renowned for luxurious cabins, spacious seating, quality entertainment systems, and fine dining experiences, Singapore Airlines prioritizes passenger comfort.
3. **Reliability and Safety:** The airline maintains an exceptional safety record, instilling trust and confidence among passengers, contributing significantly to its positive reputation.

Singapore Airlines' unwavering commitment to delivering top-notch service, ensuring passenger comfort, and upholding high safety standards has cemented its position as a global industry leader in aviation excellence.

2. High-Level Description of Solution:

Our approach involves a multi-faceted analysis combining both textual reviews and numerical ratings to gain comprehensive insights into airline customer experiences. Leveraging advanced natural language processing techniques, such as topic modeling through LLama2 and sentiment analysis via ChatGPT 3.5 Turbo, we've extracted nuanced sentiments associated with specific topics from textual reviews. This method allows us to categorize sentiments related to services like food and beverage, wifi connectivity, and cabin services, providing a deeper understanding of customer opinions.

For the numerical ratings, we have developed a machine learning-based recommendation system. This system tests diverse models such as RandomForest, K-Nearest Neighbors (KNN), Decision Trees, and Logistic Regression. Through rigorous testing and evaluation, we've created a binary recommendation system that considers various numerical attributes, such as ratings for food and beverage, wifi quality, cabin services, etc., to predict whether a customer would recommend an airline based on their experiences.

3. How it Works:

Architecture of Approach: The architecture comprises two primary components:

1. Textual Analysis using KeyBERT, LLama2 and ChatGPT 3.5 Turbo for sentiment extraction and topic modeling.
    - First the text reviews are preprocessed and cleaned and prepared for topic modelling.
    - The topic modelling process performed by KeyBERT groups the reviews into similar characteristics and accordingly we get details on these groups.
    - Llama2 takes these topics and creates written summaries of the topics which can then be easily identified by any viewer.
    - We then use ChatGPT 3.5 Turbo to categorize each group. So that if a group of reviews refers to cabin service, it will be categorized as cabin service. We have numerical ratings to validate the categorization by our algorithm.
    - Finally, the NLTK toolkit performs sentiment analysis on Llama2 outputs, giving us a sentiment score to better understand how people feel about a certain service or category.
2. Numerical Analysis utilizing diverse machine learning models for creating a recommendation system based on numerical ratings.

- Various classification models were applied on it, like Logistic Regression, KNN, RandomForest, and DecisionTree to experiment and find the most appropriate model.
- We found that with our data all the models accurately deliver on what is required.

Inner Workings of Components:

- **Textual Analysis:**

LLama2 and ChatGPT 3.5 Turbo are employed to process text reviews, extracting sentiments and associating them with specific topics. The sentiment-tagged topics provide insights into customer feelings regarding different airline services. These generative AI models come with some flexibility. Mainly we tuned the temperature variable, as well as the model nodes capacity limit. With appropriate resources, Llama2 can utilize upto 70 billion nodes. For our application we stuck with a much lesser 13 billion.

- **Numerical Analysis:** The machine learning-based recommendation system involves feature engineering using numerical ratings. Various models are tested and evaluated to ascertain their effectiveness in predicting customer recommendations.

The input data has been meticulously extracted and feature engineered to deliver best results when passed through a model. All the ratings are integers values, and while continuous do not range out of scope, and neither does our target variable.

Model Improvement Steps: We iteratively validated models, and explored different feature engineering techniques to enhance the performance of both textual sentiment extraction and the numerical recommendation system.

4. Evaluation, Outcomes, and Discussion:

Results and Evaluation Metrics: We've obtained promising results from both the textual sentiment analysis and the numerical recommendation system. The topic modelling and the corresponding sentiments are consistent with our analysis. Metrics such as accuracy, precision, recall, and F1-score were used to evaluate the performance of our models.

Text Analysis Results:

Topic	Count	Name	CustomName	Keyword	presentati	KeyBERT	Llama2	esentative_	SentimentScore
-1	1356	-1_flight_a	"American Airlines Flight Time"	Flight Time	['flight', 'ai	['airline', 'f	['"America	['horrible e	0
0	163	0_rude_at	"Airline Staff Unprofessional Behavior"	Cabin Staff Service	['rude', 'at	['airline', 'p	['"Airline S	['unprofes	0
1	148	1_custome	"Terrible Customer Service at the Gate - Worst	Ground Service	['custome	['airline', 'f	['"Terrible	['worst cus	-1
2	57	2_good_lo	"Admiral's Club Lounge - Excellent Business Cla	Food & Beverages	['good', 'lo	['inflight', '	['"Admiral'	['cabin cre	1
3	42	3_seat_together_extra_paid			['seat', 'to	['airline', 'f	['\nAirlines	['flying eve	0
4	35	4_worst_f	"Worst Airline Ever"	Overall Ratings	['worst', 'fl	['airline', 'f	['"Worst A	['worst air	-1
5	29	5_charlott	"Charlotte AA Refund Car Cancelled Mile Airport Flight Myrtle Would".		['charlotte	['airline', 'f	['"Charlott	['seen serv	0
6	28	6_disappo	"Disappointing Travel Experience in Chicago"	Overall Ratings	['disappoi	['airline', 'c	['"Disappo	['disappoi	-0.6
7	24	7_miami_t	"Miami International Airport Tire Pilot Schedule	Ground Service	['miami', 't	['flight', 'm	['"Miami In	['paid alm	0
8	24	8_good_pl	"A Pleasant and Comfortable Experience in Ha	Overall Ratings	['good', 'pl	['inflight', '	['"A Pleasa	['pleasantl	0.566666667
9	23	9_dallas_v	"Connecting Flight in Dallas After Waiting All M	Ground Service	['dallas', 'w	['airline', 'f	['"Connect	['nightmar	0
10	22	10_child_f	"Baby on Board: A Family's Journey with a Sma	Overall Ratings	['child', 'fa	['airline', 'p	['"Baby on	['disappoi	-0.25
11	21	11_charlot	"Worst flight experience ever connecting in Ch	Overall Ratings	['charlotte	['airline', 'f	['"Worst fl	['worst air	-1

Recommendation System Results:

ML Model	Precision	Recall	F1- score
Logistic Regression	0.966	0.967	0.967
K Nearest Neighbors	0.964	0.964	0.964
Decision Trees	0.951	0.951	0.951
Random forest	0.964	0.964	0.964

Discussion and Takeaways: Our analysis reveals valuable insights into customer sentiments regarding specific airline services. The sentiment-categorized topics shed light on areas of strength and improvement for airlines. Additionally, our recommendation system showcases the effectiveness of machine learning models in predicting customer recommendations based on numerical ratings and inputs.

We observed that the combined approach of textual sentiment analysis and numerical rating-based recommendation systems provides a holistic view of customer experiences. While our models perform well, there may be scenarios where external factors or outliers impact predictions. Continuous refinement and adaptation of models are essential for robust recommendations and actionable insights for airlines. Revision is required especially when standardization across different sources of data are considered.

In summary, our approach amalgamates cutting-edge NLP techniques and machine learning models to comprehensively analyze and understand airline customer experiences, offering airlines valuable insights for enhancing services and customer satisfaction. If the airline delivers on what the models recommend, the airline should receive popularity and respect like in the case of Singapore Airlines as mentioned earlier.

References:

1. Touvron, Hugo, et al. "Llama 2: Open foundation and fine-tuned chat models." *arXiv preprint arXiv:2307.09288* (2023)
2. Shin, Hunsik, Hye Jin Lee, and Sungzoon Cho. "General-use unsupervised keyword extraction model for keyword analysis." *Expert Systems with Applications* 233 (2023): 120889.
3. Kumar, S., Zymbler, M. A machine learning approach to analyze customer satisfaction from airline tweets. J Big Data 6, 62 (2019). https://doi.org/10.1186/s40537-019-0224-1
4. Novaković, Jasmina Dj, et al. "Evaluation of classification models in machine learning." *Theory and Applications of Mathematics & Computer Science* 7.1 (2017): 39.
5. Al-Qahtani, Rana, and Princess Nourah bint Abdulrahman. "Predict sentiment of airline tweets using ML models." (No. 5228). EasyChair (2021).