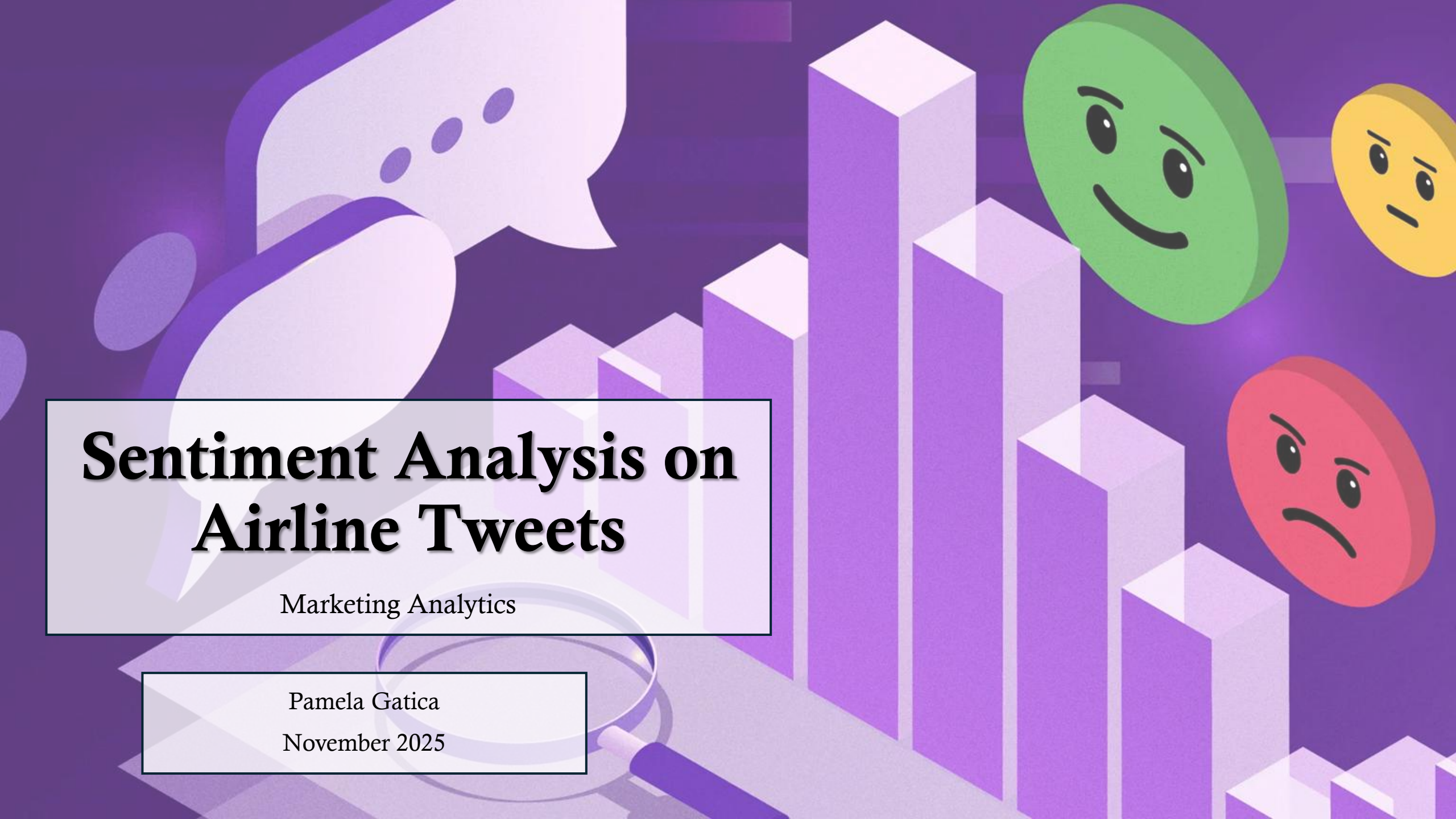


Sentiment Analysis on Airline Tweets

Marketing Analytics

Pamela Gatica

November 2025

Introduction

- Sentiment Analysis is a Natural Language Processing (NLP) technique used to **automatically detect emotions or opinions** expressed in text.
- It helps classify whether a piece of text is:
  - ✓ **Positive**
  - ✓ **Neutral**
  - ✓ **Negative**



Dataset Description



The dataset contains **14,640 customer tweets** about airline experiences.

The original dataset included **15 columns; 3 columns were removed** because of large number of missing values.

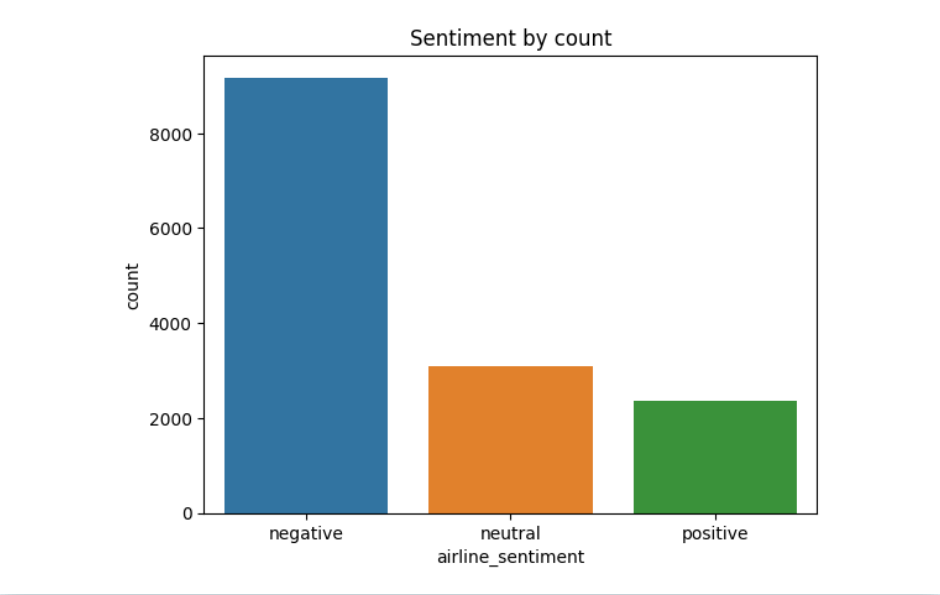
Key columns used in the project:

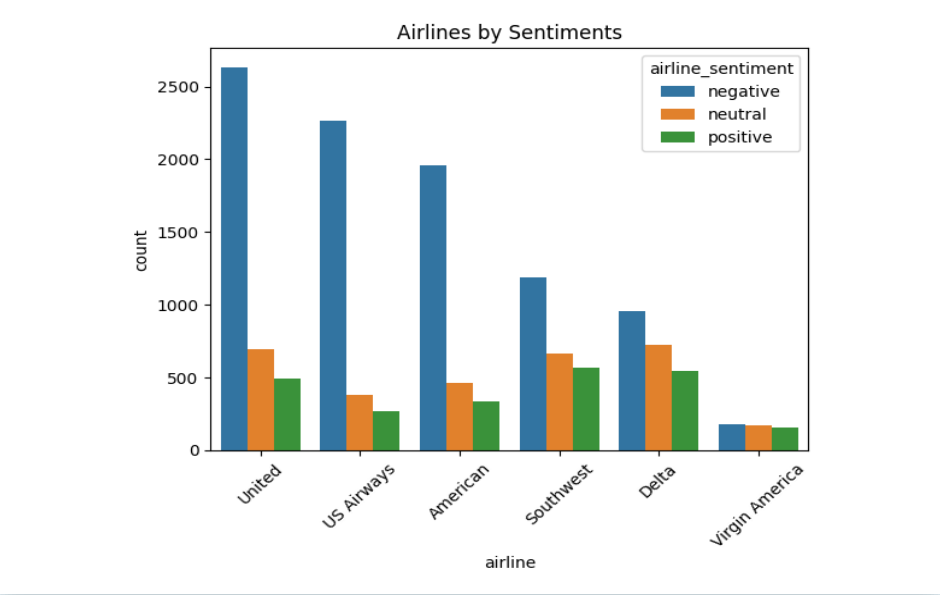
text (raw text)

airline_sentiment (positive, neutral, negative)

Sentiment Distribution

- **Negative tweets dominate the dataset**, representing the majority of customer feedback.
- **Positive tweets are the least common.**
- United Airlines received the highest number of negative tweets.
- Virgin America has noticeably fewer negative posts.
- **Positive tweets are consistently lower** across all airlines.

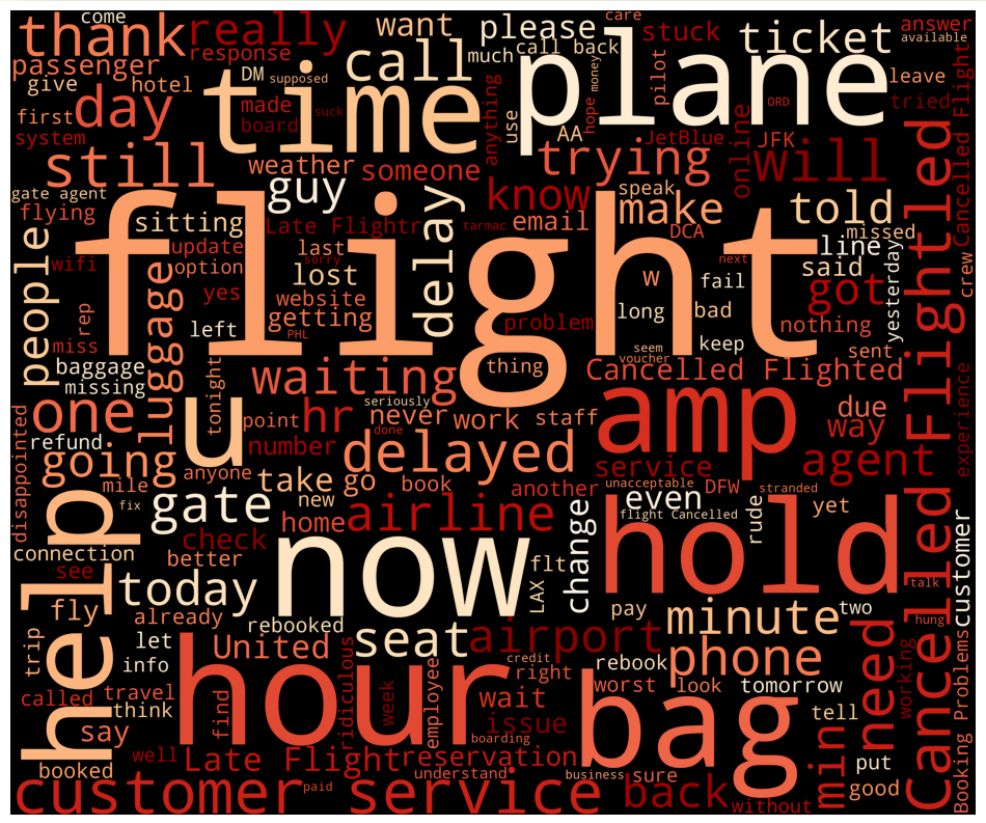




Word Clouds

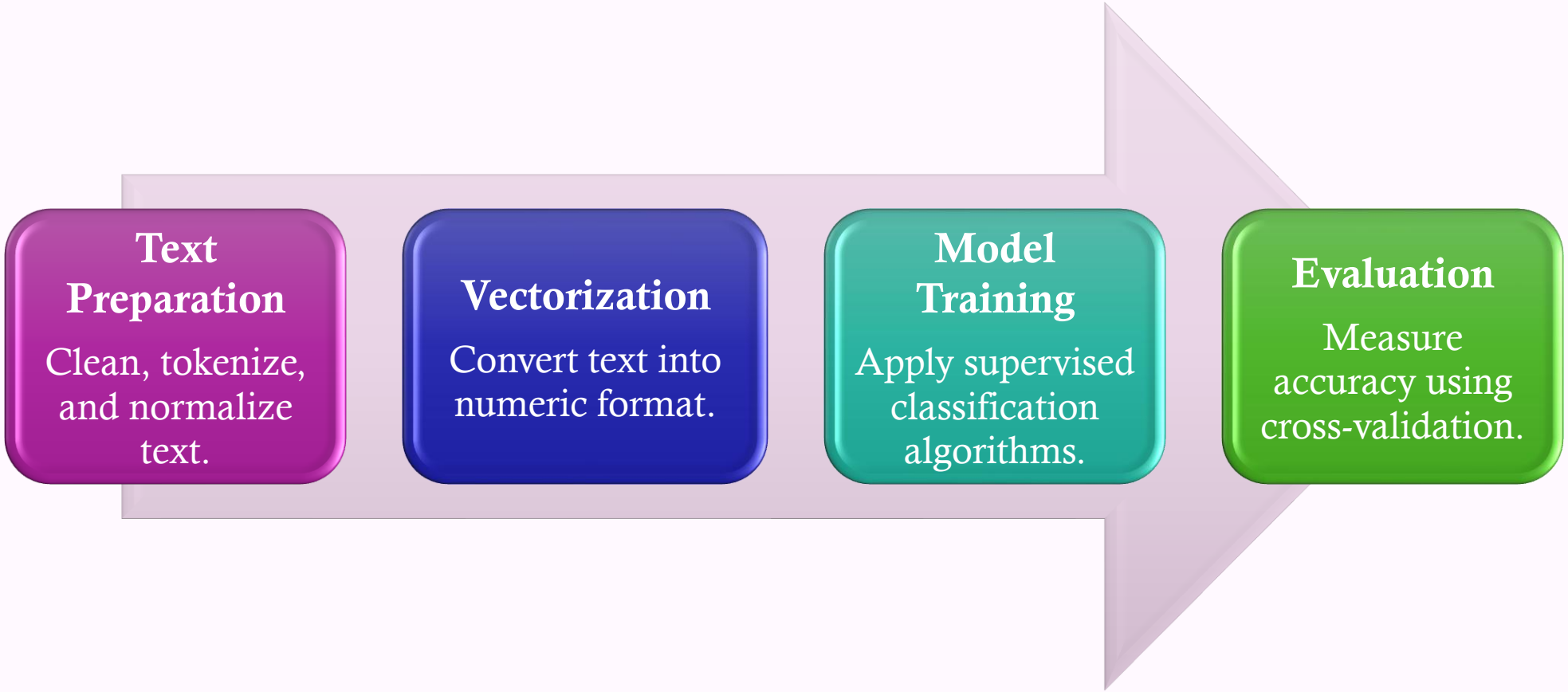
Positive vs Negative Sentiment





Machine Learning for Sentiment Classification

Used to automatically classify tweets into **positive**, **neutral**, or **negative** sentiment.



Model Results and Insights

Model	Count Vectorizer	TF-IDF Vectorizer
Logistic Regression	78.921%	76.858%
Random Forest	76.926%	76.660%
Support Vector Machine (SVM)	77.152%	76.701%

- ✓ **Count Vectorizer delivered higher accuracy** than TF-IDF for this dataset.
- ✓ **Logistic Regression achieved the best overall performance.**
- ✓ All models produced consistent results around the 76–79% range.

Conclusions

- Negative sentiment dominates the dataset.
- Delays, cancellations, and customer service issues are the biggest pain points.
- Logistic Regression provided the highest accuracy for classification.

Recommendations

- Improve communication during delays and cancellations to reduce frustration.
- Strengthen customer service responsiveness, especially during peak travel periods.
- Invest in better baggage handling processes to reduce lost-luggage incidents.




THANK YOU