Title: *Predicting Airline Tweet Sentiment Using Machine Learning and NLP*
Author: Neelam Pandey
Program: Data Science Career Track – Capstone 3 Project
Date: January 2026

1. Introduction

Airlines receive thousands of customer tweets every day, ranging from praise to severe complaints. Understanding this sentiment in real time is essential for customer experience teams, crisis management, and brand monitoring.

This project builds a machine-learning system that classifies airline-related tweets into **positive**, **neutral**, or **negative** sentiment. The goal is to help airline stakeholders quickly identify emerging issues, prioritize customer support, and track brand perception over time.

2. Problem Statement

Airlines struggle to manually monitor the volume and velocity of social media feedback. Without automated sentiment classification, teams miss early warning signs of operational issues (delays, cancellations, customer service failures).

Objective:
Develop a supervised machine-learning model that predicts tweet sentiment with high accuracy using natural language processing (NLP).

3. Dataset Overview

Source: Twitter US Airline Sentiment Dataset (cleaned version).
Rows: ~14,000 tweets
Key columns:

text

clean_text

processed_text

airline_sentiment

negativereason

tweet_created

The dataset is appropriate for the problem because it contains labeled sentiment classes and rich text features.

4. Data Wrangling

All DSM Step 2–3 wrangling requirements were completed:

- Missing values handled (negativereason, processed_text)
- Empty strings replaced with fallback text
- Critical rows with missing text removed
- Target variable encoded using LabelEncoder
- Cleaned dataset saved as twitter_airline_cleaned_for_modeling.csv

5. Exploratory Data Analysis (EDA)

Three core visualizations were selected for the final report:

Figure 1 — Sentiment Distribution

Shows class imbalance (majority negative tweets).
Supports the need for stratified sampling.

Figure 2 — Most Common Words in Negative Tweets

Highlights operational pain points (e.g., “delay”, “customer service”, “late”).

Figure 3 — Tweet Volume Over Time

Shows spikes in negative sentiment during operational disruptions.

These figures support the narrative that airlines face recurring customer-experience issues that can be monitored via NLP.

6. Preprocessing & Feature Engineering

Rubric-aligned preprocessing steps:

- **Dummy features:** TF-IDF vectorization (1–2 grams, 5000 features)
- **Magnitude standardization:** TF-IDF inherently normalizes feature magnitude
- **Train/test split:** 80/20 with stratification
- **Label encoding:** Converts sentiment to numeric classes

7. Modeling Approach

Four models were trained:

1. Logistic Regression
2. Linear SVC
3. Random Forest
4. Gradient Boosting

Each model used the same TF-IDF pipeline for fairness.

8. Model Performance Comparison

Linear SVC achieved the highest accuracy and strongest classification report across all classes.

9. Final Model Selection

Chosen Model: Linear SVC
Reason:

- Best accuracy
- Best precision/recall for minority classes

Fast training and inference

- Robust for high-dimensional sparse text data

The model was saved as:
best_twitter_airline_model_LinearSVC.joblib

10. Results & Interpretation

The model correctly identifies negative sentiment with high precision, enabling airlines to:

- Detect customer dissatisfaction early
- Prioritize support tickets
- Monitor operational issues in real time

Confusion matrix analysis shows strong separation between positive and negative classes, with some confusion between neutral and negative tweets—expected due to linguistic ambiguity.

11. Recommendations for Stakeholders

1. Real-Time Monitoring Dashboard

Deploy the model into a dashboard that alerts teams when negative sentiment spikes.

2. Root-Cause Analysis

Use the negativereason field to categorize complaints (delays, cancellations, customer service).

3. Customer Service Automation

Integrate the model into a triage system that routes high-severity tweets to human agents.

12. Future Work

- Add hyperparameter tuning (GridSearchCV)
- Incorporate deep learning models (BERT, RoBERTa)
- Expand dataset to include multiple airlines and international tweets
- Add topic modeling for deeper insights

13. Conclusion

This project demonstrates that machine-learning models can accurately classify airline tweet sentiment and provide actionable insights for customer-experience teams. The final Linear SVC model delivers strong performance and is ready for deployment in a real-world monitoring system.