SUMMARIZATION OF AIRLINE CUSTOMER REVIEWS USING NLP TECHNIQUES

ANLT-293-ON1 Spring 2025

Hsin-Yu Liao   Chun-Yen Lin

TABLE OF CONTENTS

INTRODUCTION

In the airline industry, online reviews reflect customer experiences and expectations. However, their unstructured and large-scale nature makes manual analysis difficult, limiting timely insights for both travelers and airlines.

This project uses NLP techniques to:

- Classify each sentence into one of 17 service categories.
- Analyze sentiment polarity to find key emotional highlights.
- Combines zero-shot classification (BART) and sentiment scoring (RoBERTa)
- Delivers concise summaries that reflect both topics and emotions.

BACKGROUND

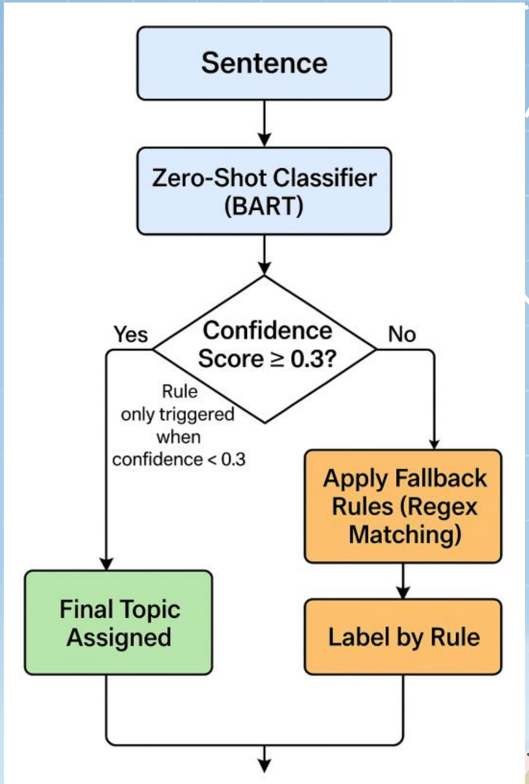
- Customer reviews have become a critical touchpoint for travelers and airlines.
- Reviews offer valuable insights but are often lengthy, subjective, and cover multiple topics.
- Manual analysis is inefficient, especially at scale.
- Automating this process transforms raw text into structured insights, improving accessibility and usability of feedback data.

METHODOLOGY - DATA PREPARATION

- Dataset from Kaggle: 3,700 British Airways reviews
- Focused on the `ReviewBody` column
- Removed missing values and short reviews (<10 characters)
- Applied sentence segmentation using ***spaCy*** with `en_core_web_md`
- Converted all text to lowercase and removed special characters
- Produced clean, sentence-level input for analysis

METHODOLOGY - TOPIC CLASSIFICATION

- Used `facebook/bart-large-mnli` for zero-shot topic classification
- Compared each sentence against 17 predefined airline-related labels
- Fallback used keyword-based regex patterns (e.g., “screen” → In-flight Entertainment)



METHODOLOGY - SENTIMENT ANALYSIS

- Used `cardiffnlp/twitter-roberta-base-sentiment`
- Categories: Positive, Neutral, Negative
- Sentiment output converted to polarity scores: +1 to -1
- Neutral = 0, Positive > 0, Negative < 0
- Enabled comparison of emotional intensity across sentences

METHODOLOGY - SUMMARIZATION STRATEGY

Most emotional sentence

- highest absolute polarity score
- Preference given to “Overall Airline Experience” or “Value for Money”

Three additional topic sentences

- high confidence (≥ 0.25)
- Ensured topic diversity by avoiding repeated labels
- Produced concise summaries reflecting content and tone

RESULT - REVIEW 1: POSITIVE EXPERIENCE

- Reviewer praised seat, crew, food, and entertainment
- Emotional sentence: “BA can be among the world’s best airlines” (score: 0.97)
- Topic sentences included:
  - In-flight Entertainment
  - In-flight Food and Drinks
  - Staff & Service Attitude

Aspect Category	Representative Sentence
In-flight Entertainment	In-flight entertainment offered a great selection of films and audio.
In-flight Food and Drinks	On board food for dinner and breakfast was good and there was a well chosen selection of wines.
Staff & Service Attitude	The crew were attentive, friendly and very professional.
Most Emotionally Charged Sentence	This flight showed that BA can be among the world’s best airlines. Sentiment Score: 0.97

Table 1 Model-Generated Review Summary for Review 1

RESULT - REVIEW 2: NEGATIVE EXPERIENCE

Aspect Category	Representative Sentence	Confidence Score	Sentiment Score
Overall Airline Experience	The aircraft is very old, cabin configuration is very old and tired.	0.22	-0.93
In-flight Entertainment	IFE screens have not been changed since they were first installed.	0.61	0
In-flight Entertainment	My iPod has a larger and more responsive screen. *	0.55	0.9
In-flight Food and Drinks	After take off, drinks were offered, followed by a hot meal.	0.58	0
In-flight Food and Drinks	Food choices ran out in the first row. *	0.55	-0.63
Seat Comfort	Seats were uncomfortable, footrests were jammed. *	0.63	-0.95
In-flight Entertainment	Poor movie choices, miniature screen and uncomfortable seats.	0.42	-0.94
Value for Money	Having flown Norwegian on their B787 in their premium cabin on the same route, BA is a waste of my money. *	0.56	-0.82

Table 2 Selected Rows from Sentence Classification Output for Review 2

RESULT - REVIEW 2: NEGATIVE EXPERIENCE

- Only one summary aspect included due to filtering
- Fallback rules used in 4+ sentences
- Topic sentences included :
  - Seat Comfort
  - In-flight Entertainment
  - In-flight Food and Drinks

Aspect Category	Representative Sentence
Seat Comfort	Seats were uncomfortable, footrests were jammed.
In-flight Entertainment	IFE screens have not been changed since they were first installed.
In-flight Food and Drinks	After take off, drinks were offered, followed by a hot meal.
Most Emotionally Charged Sentence	The aircraft is very old, cabin configuration is very old and tired. Sentiment Score: -0.93

Table 3 Model-Generated Review Summary for Review 2

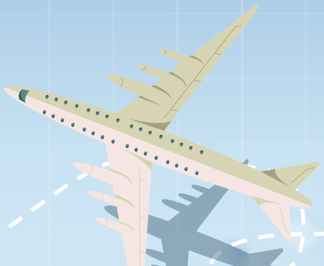
FUTURE WORK

Improve Classification Accuracy

- Handle class imbalance using SMOTE
- Apply hyperparameter tuning
- Introduce confidence scoring for better interpretability

Real-World Deployment

- Develop a browser-based application (e.g., Chrome extension)
- Display real-time summaries & sentiment insights on airline review sites
- Reduce the need to read lengthy texts
- Deliver instant, user-friendly insights



CONCLUSION

From unstructured reviews to actionable insights through hybrid NLP.

- Used NLP to classify and summarize airline reviews
- Combined zero-shot models with rule-based fallback for better accuracy
- Added sentiment analysis to highlight emotional tone
- Sentence-level processing enabled precise multi-topic understanding

REFERENCES

Anshul Chaudhary, & Muskan Risinghani. (2023). *Airline reviews*. Kaggle. https://doi.org/10.34740/KAGGLE/DS/4044107

Cardiff NLP. (2021). *Twitter-roBERTa-base for Sentiment Analysis*. Hugging Face. https://huggingface.co/cardiffnlp/twitter-roberta-base-sentiment

Honnibal, M., & Montani, I. (2021). *spaCy 3: Industrial-strength Natural Language Processing in Python* [Software]. Explosion. https://doi.org/10.5281/zenodo.1212303

REFERENCES

Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., ... & Zettlemoyer, L. (2020). *BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension*. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (pp. 7871–7880). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.acl-main.703

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., ... & Rush, A. M. (2020). *Transformers: State-of-the-art natural language processing*. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, 38–45. https://doi.org/10.18653/v1/2020.emnlp-demos.6

THANKS!

DO YOU HAVE ANY QUESTIONS?