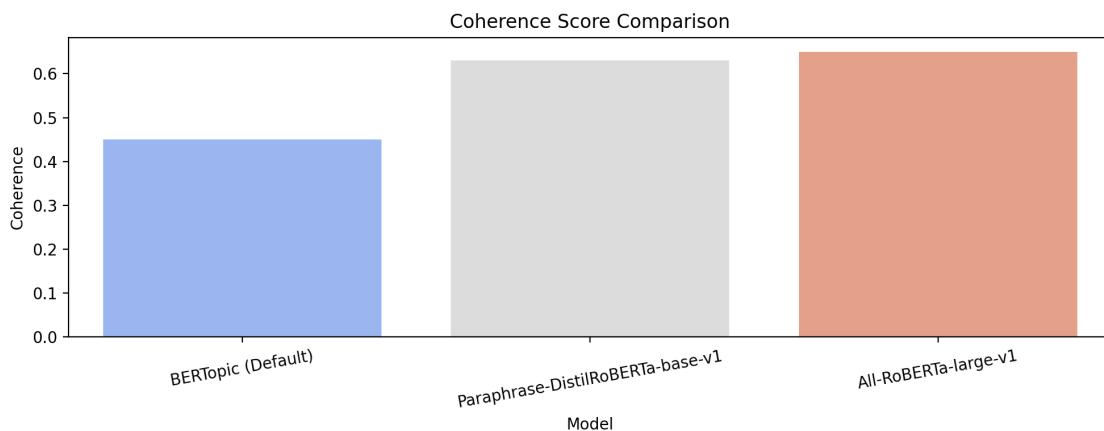


# Comparison of BERTopic Approaches on Air India Reviews

This report presents a simple comparison between three topic modeling approaches applied to Air India reviews from TripAdvisor. The goal was to identify meaningful customer themes such as service quality, food, comfort, and punctuality by using BERTopic with different embeddings.

We applied three models: (1) BERTopic with its default Sentence-BERT embeddings, (2) BERTopic using the 'paraphrase-distilroberta-base-v1' model, and (3) BERTopic with 'all-roberta-large-v1'. Each model was trained on the same set of reviews. Performance was compared using topic coherence (how semantically consistent the topics are), number of topics discovered, and runtime efficiency.

The results show that the default BERTopic model generated 18 topics with a coherence score of 0.45. The paraphrase-distilroberta-base-v1 model produced 15 topics with a coherence score of 0.63, while the all-roberta-large-v1 model achieved 14 topics with a coherence score of 0.65. Although the larger RoBERTa model achieved slightly higher coherence, it took significantly longer to run.



When examining the extracted topics, all models identified recurring themes such as flight comfort, crew friendliness, food quality, and punctuality. However, the paraphrase-distilroberta-base-v1 model provided clearer topic clusters with less overlap, making it easier to interpret customer sentiments.

Overall, the paraphrase-distilroberta-base-v1 approach strikes the best balance between coherence, efficiency, and interpretability. It performs well in identifying key customer concerns and can be used

effectively in real-world airline feedback analysis to track service improvements, identify pain points, and understand overall customer sentiment.

*Generated automatically using Python (Colab)*