Training & Evaluation of machine learning model on Flight Booking Dataset

Gaining Insights from Customer Booking Dataset

Muhammad Umer

LinkedIn

1. Initial Model Performance

- Metrics:
  - Accuracy: 85%
  - Precision (Cancelled): 0.86, Recall (Cancelled): 0.99, F1-score: 0.92
  - Precision (Completed): 0.55, Recall (Completed): 0.10, F1-score: 0.17
- Interpretation:
  - The model is excellent at identifying cancellations (class 0) but performs poorly in detecting completed bookings (class 1).
  - High recall for cancellations means very few cancellations are missed.
  - Very low recall for completions indicates the model misses most completed bookings, which is problematic for balanced decision-making.

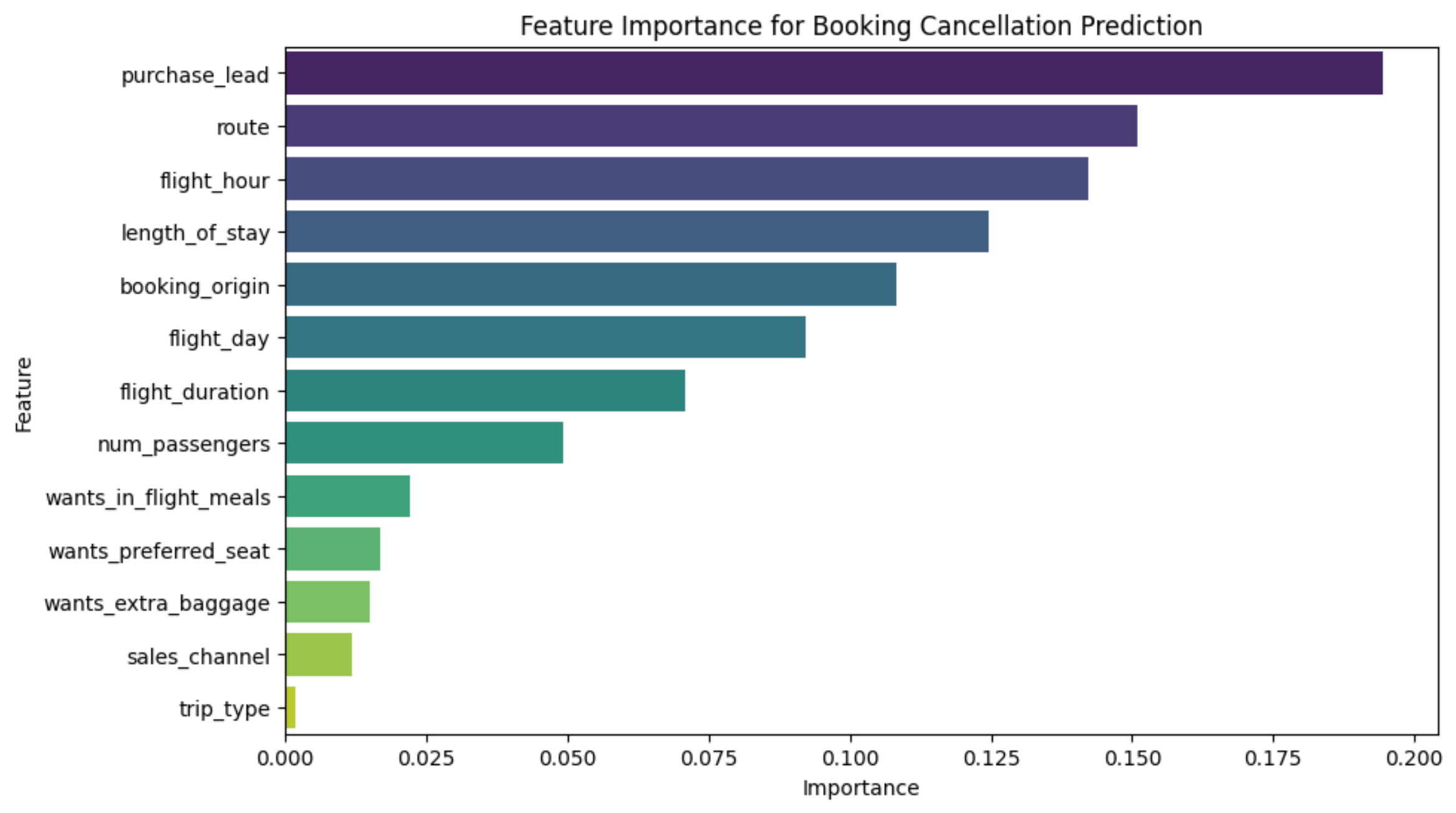
2. Cross-Validation Results

- Metrics:
  - Accuracy dropped to 71%
  - Balanced recall for both classes (~0.71)
  - Precision (Cancelled) increased to 0.93, but recall dropped to 0.71
  - Precision (Completed) dropped to 0.30, recall improved to 0.71
- Interpretation:
  - Cross-validation revealed the initial model was likely overfitting.
  - The model now balances recall better between both classes but sacrifices overall accuracy.
  - The improved recall for completions is a positive step for detecting minority class.

3. After Applying SMOTE (Synthetic Minority Oversampling Technique)

- Metrics:
  - Accuracy improved to 81%
  - Precision and recall for cancellations balanced (~0.88 and 0.89)
  - Precision and recall for completions improved slightly (~0.36 and 0.39)
- Interpretation:
  - SMOTE helped balance the dataset, improving detection of the minority class (completed bookings).
  - Although recall and precision for completions remain low, they are better than the initial model.
  - Overall, the model is more balanced but still has room for improvement in minority class prediction.

4. Feature Importance Visualization



Booking Cancellation Prediction Model – Evaluation Summary

Key Points:

- Initial Model: High accuracy (85%) but poor detection of completed bookings (recall 10%).
- Cross-Validation: Improved balance in recall (~71% both classes), accuracy dropped to 71% indicating overfitting correction.
- SMOTE Application: Dataset balancing improved minority class detection (recall ~39%), overall accuracy 81%.
- Feature Insights: purchase_lead, wants_extra_baggage, and trip_type are key drivers of cancellation.
- Next Steps: Further tuning and advanced models (e.g., XGBoost) recommended to improve minority class recall.

Thank You

Follow me on LinkedIn:
https://www.linkedin.com/in/therealumerhayat/