Summit Data Group

Fall 2024 Case Competition

info@sdg.com
www.summitDataGroup.com

Summit Data Group

Introduction

Overview

Ocean Crest Hotel faces revenue loss and disruptions due to high booking cancellations, needing a data-driven solution to predict and reduce them.

Business Problem

Ocean Crest Hotel faces a 37% booking cancellation rate, disrupting revenue forecasting, resource planning, and operational efficiency.

Solution

Using predictive analytics to improve processes, optimize distribution, and enhance customer management.

Summit Data Group

Key Findings

N Non-Refundable Deposits (Highest Impact)

O Online marketing channel

A Average Daily Rate (ADR)

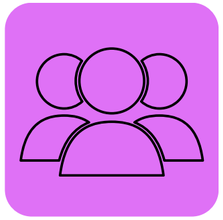
P Cancellations: Previous and Non-Cancellations

L Lead Time – Short and Long

CR Customer Requests: Special Requests, Parking requests etc.

Summit Data Group

Recommendations



Loyalty Program 2.0

Tiered Deposit System

Dynamic Pricing Strategy

OTA Booking Enhancement

Lead Time

Summit Data Group

Business Improvement

01

Revenue

Projected 28% year-over-year increase

02

Operations

Improved efficiency and resource management

03

Guest experience

Enhanced satisfaction and personalization

04

Distribution

Reduced dependence on high-cancellation platforms

05

Cancellation Rate

A reduction in the cancellation by 15% in the first year

Summit Data Group

Implementation Roadmap

Short-term (0-3 months)

- Launch a tiered deposit system

Medium-term (3-6 months)

- Introduce dynamic pricing strategy

Long-term (6-12 months)

- Optimize OTA partnerships and direct booking channels

Summit Data Group

Conclusion

- A data-driven approach to tackle cancellation challenge
- Holistic strategy addressing multiple factors
- Potential for significant revenue growth and operational improvement

Summit Data Group

What’s Next?

- Prioritize recommendations
- Allocate resources
- Set implementation timeline

Summit Data Group

The Engine Driving the Solution

Summit Data Group

Data Overview & Preparation

Data Overview

- **Total Observations:** 112,941 rows
- **Total Features:** 31 columns
- **Period**: Data covers bookings from 2015 to 2017.

Key Column

- **Is Canceled - Response**
- **ADR** (Average Daily Rate)
- **Lead Time**
- **Special Request**
- **Booking Channel**
- **Country**

Data Preparation

Data Cleaning

- Handling of Missing Values
- Handling of Outliers
- Improving data categorization

Feature Engineering

- Geographical Aggregation
- Variable Aggregation
- Dimensionality Reduction
- New Variables

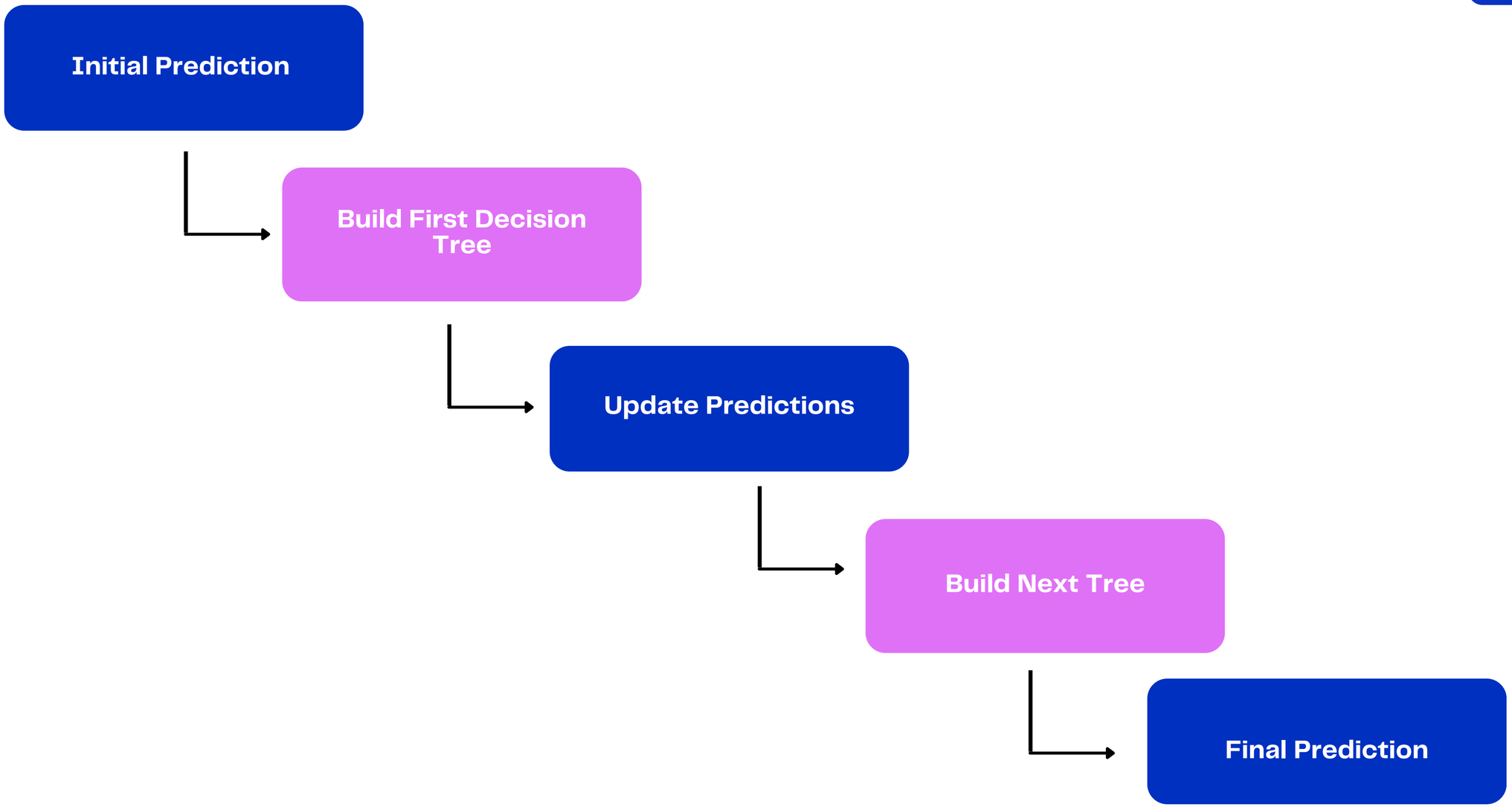
Summit Data Group

Model choice

Models	Training AUC	Testing AUC	Pro	Cons
Decision Tree	0.660	0.73	Interpretable, handles non-linear relationships	Prone to overfitting
Random Forest	0.8980	0..825	Good performance, handles feature interactions	Less interpretable than single trees
XGBoost	0.9260	0..925	Best performance, handles complex relationships	Requires careful tuning, less interpretable
LightGBM	0.9000	0.825	Fast training, handles large datasets	Slightly lower performance than XGBoost

Summit Data Group

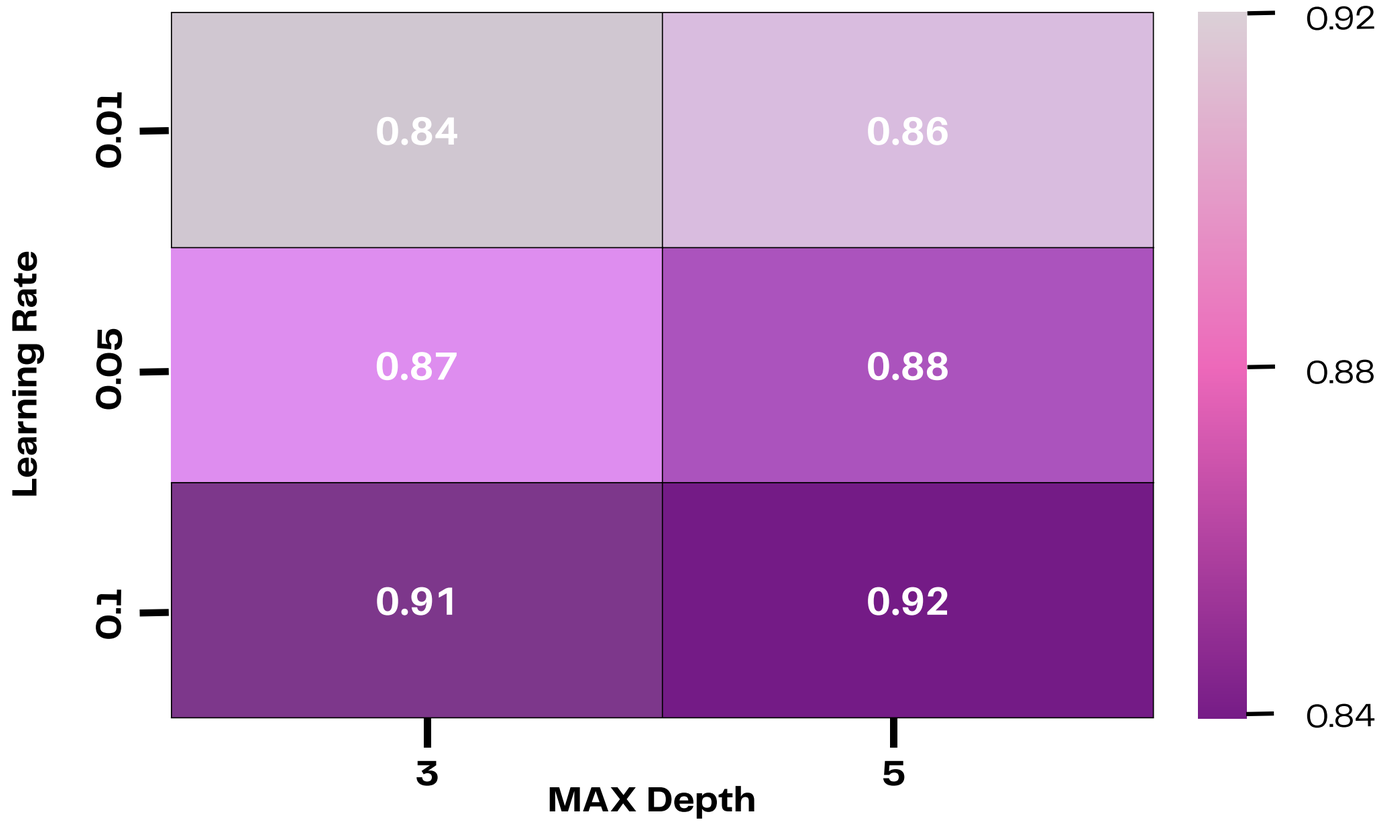
How XGBoost Works



Summit Data Group

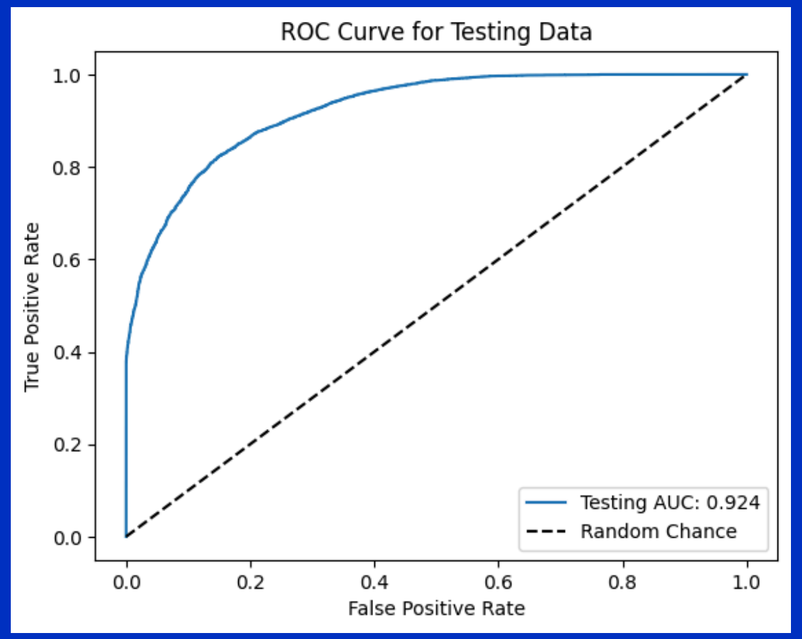
Avoiding Overfitting

- Use of ensemble methods
- Cross-validation
- Feature Engineering and Selection
- XGBoost In-Built Regularization

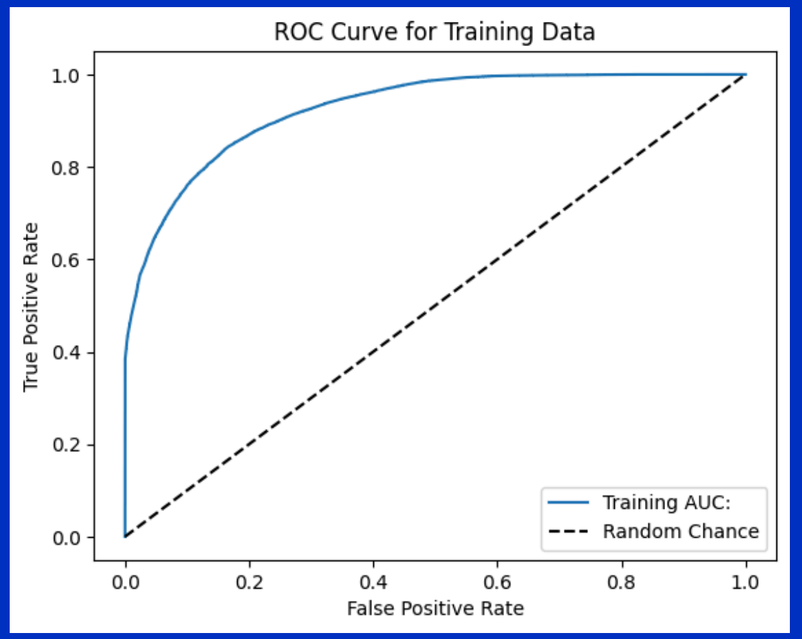


Summit Data Group

Model Evaluation – AUC Scores



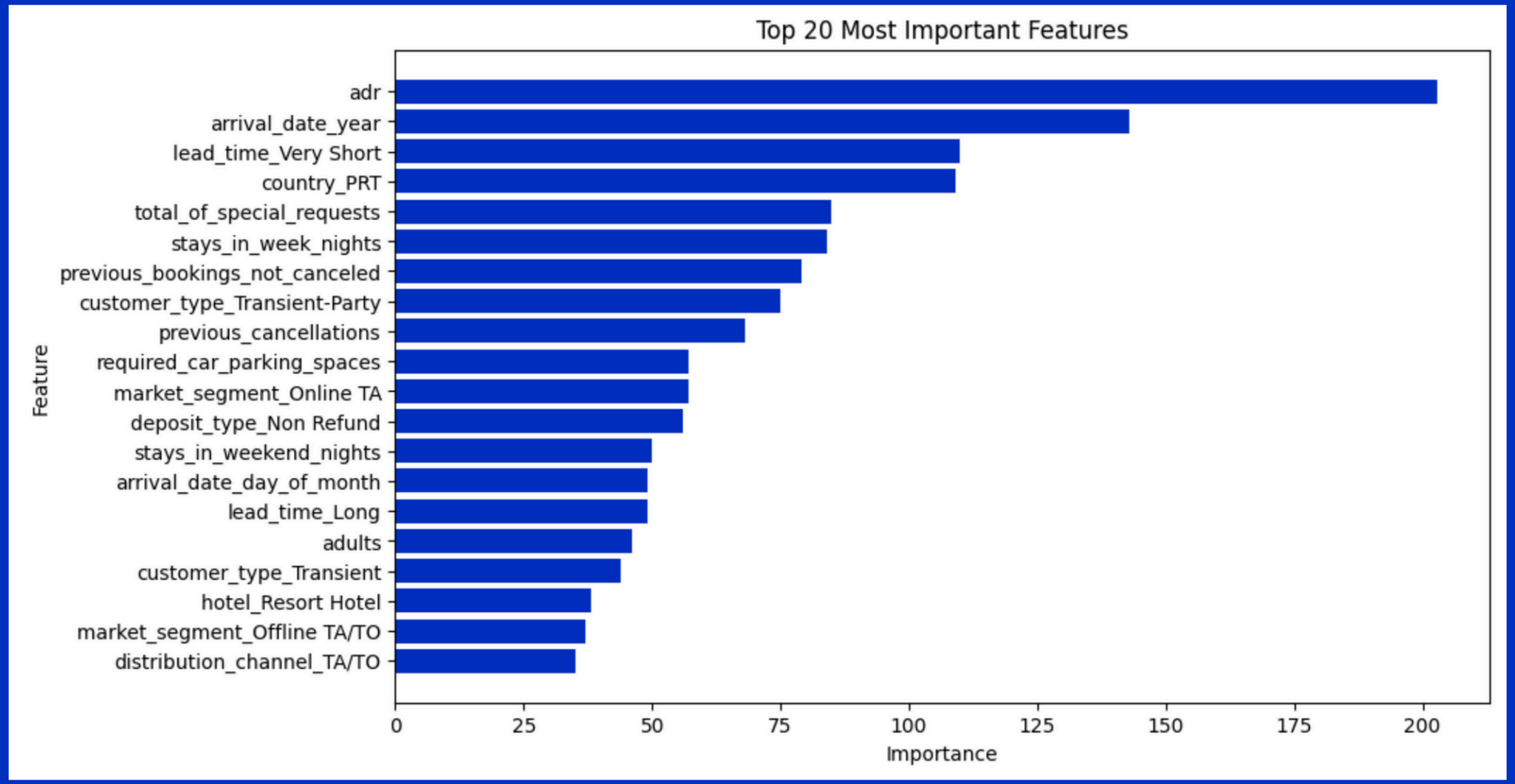
AUC for Testing Data: 0.924



AUC for Training Data: 0.926

Summit Data Group

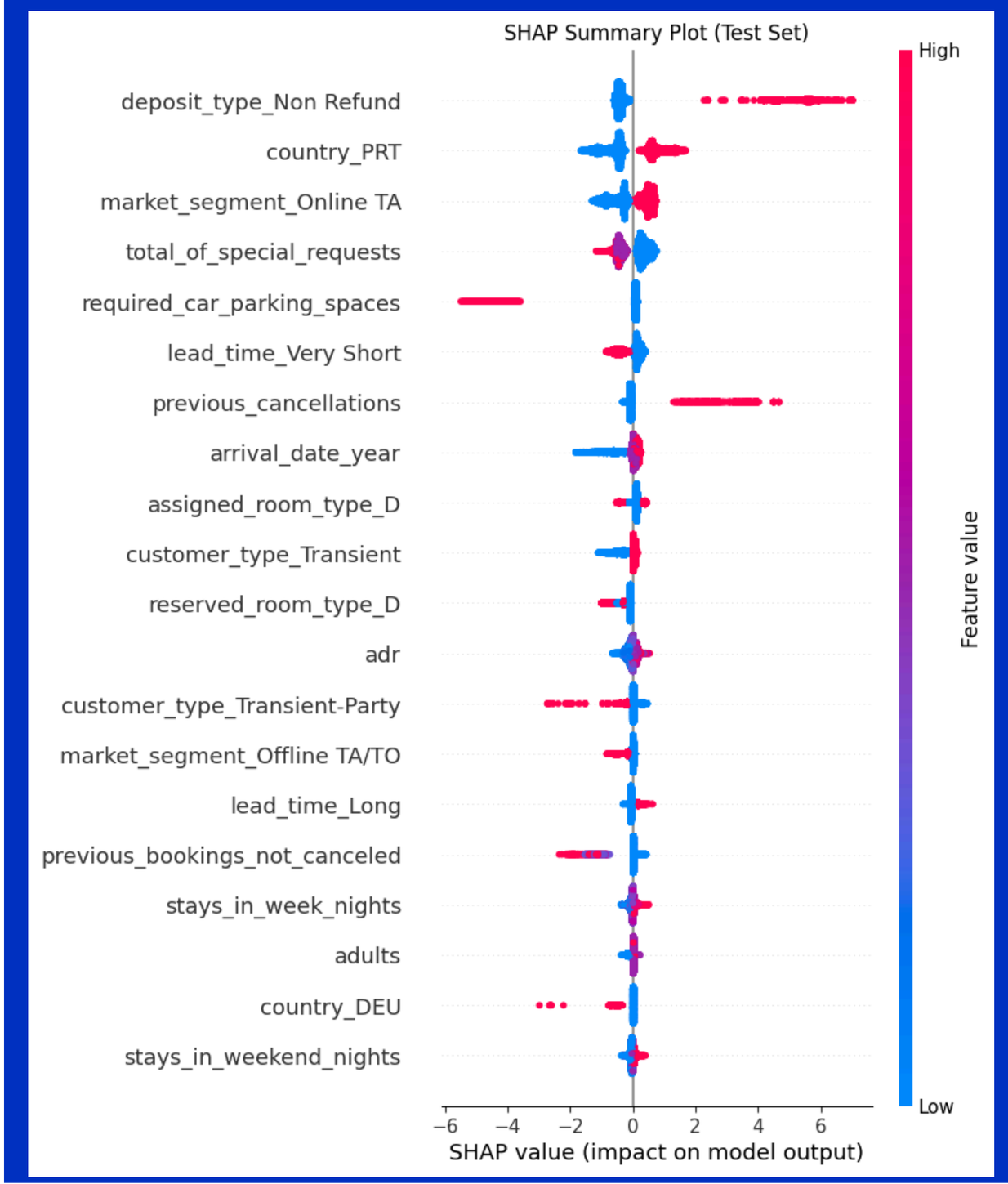
Feature Importance



Summit Data Group

SHAP Value Analysis

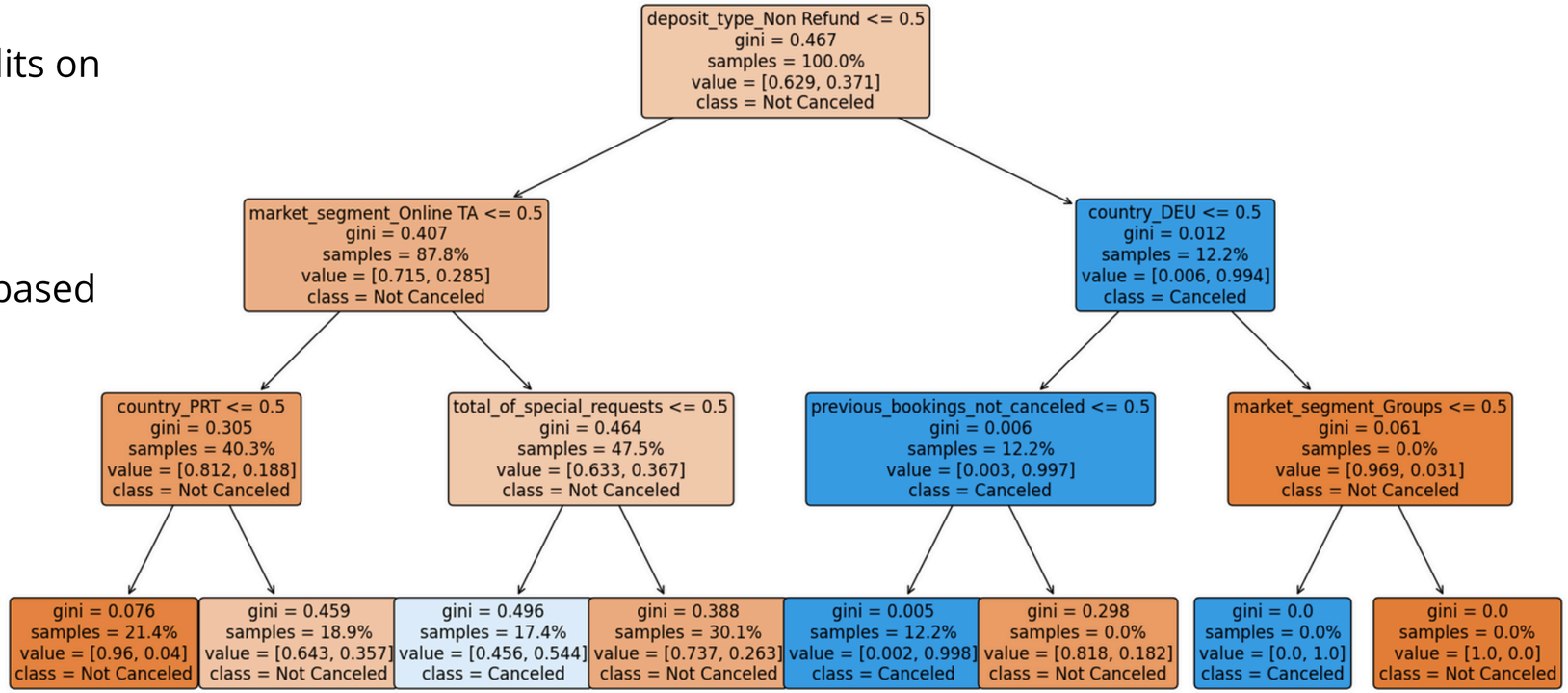
- SHAP values provide insight into how individual predictions are made.
- ADR and Lead Time are the most influential in pushing bookings toward cancellation or non-cancellation.
- SHAP analysis ensures transparency in model decisions.



Summit Data Group

Decision Tree Visualization

- A decision tree in the XGBoost model splits on features like ADR and Booking Source to predict cancellation.
- Each split further refines the prediction based on the most important factors.



Summit Data Group

Question?

- +123-456-7890
- hello@reallygreatsite.com
- www.reallygreatsite.com