Machine Learning Model for Predicting Booking Cancellations

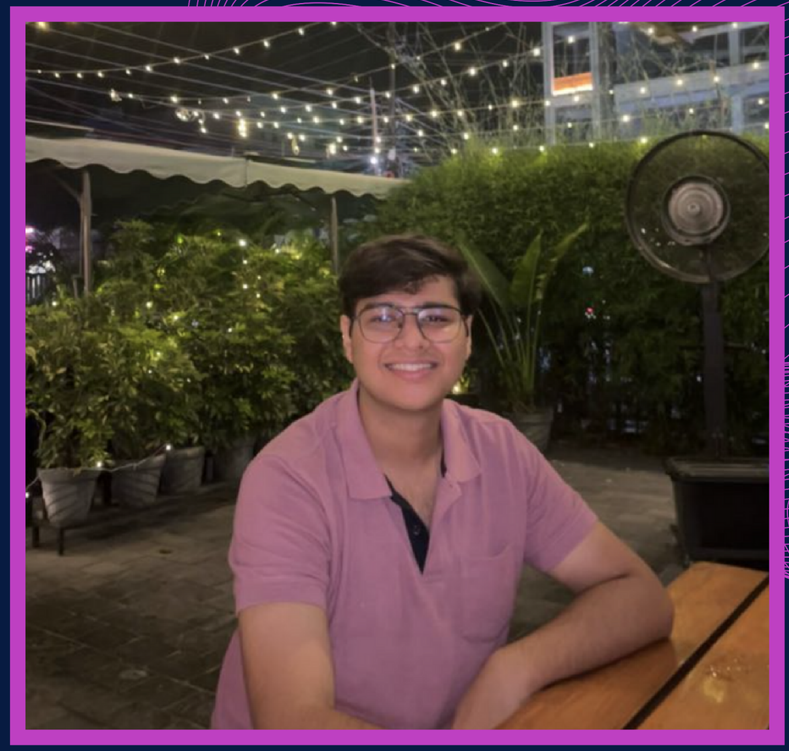
Presented by: Anmol Singhal

Internship Mode: Online

Company: Pickl.AI

Date: 28th July, 2023

Introduction



Hello everyone,

I am Anmol Singhal, a data scientist intern at Pickl.AI. Today, I am excited to share my journey and the outcomes of my project during my internship. Our goal was to develop a machine-learning model that predicts the likelihood of a booking getting canceled.

About Pickl.AI

Pickl.AI

Pickl.AI:

A pioneer in bridging the gap between theoretical and practical data science training.

Mission:

To equip young professionals with real-world problem-solving skills using data science.

Approach:

An accelerated two-month data science training program with a focus on hands-on experience.

Internship Journey

The four-week intensive internship course covered Python, Statistics, Machine Learning, Linear Regression, Classification, Ensemble Methods, Unsupervised Learning, Tableau, and SQL.
Practical Focus: Emphasis on real-world problem-solving using data science tools and techniques.

Project Overview

Project Title: Developing a Machine Learning Model to Predict Booking Cancellations.

Dataset: Provided by Pickl.AI, containing relevant booking information.

Objective: To build a predictive model that helps businesses identify potential booking cancellations in advance.

Data Exploration and Preprocessing

Missing Values:

Handled by imputing with mode and median for categorical and numerical features, respectively.

Outliers:

Detected using z-scores and removed to ensure data quality.

Feature Engineering:

Created new features and dropped irrelevant columns.

Feature Selection and Scaling

Feature Selection: Used SelectKBest with chi-squared and PCA for dimensionality reduction.

Feature Scaling: Applied MinMaxScaler to bring all features to a uniform scale.

Model Selection

Models Considered:

Logistic Regression, Random Forest, Support Vector Machine (SVM), and Gradient Boosting.

Rationale:

Chose a diverse set of models to identify the best performer.

Model Training and Evaluation

Model Evaluation Metrics:

Accuracy, Precision, Recall, and F1-score.

SVM Metrics:

Accuracy: 1.0, Precision: 1.0, Recall: 1.0, F1-score: 1.0.

Addressing Overfitting and Regularization

Overfitting:

Employed cross-validation and hyperparameter tuning to address overfitting issues.

Regularization:

Fine-tuned the models to achieve optimal generalization.

Conclusion

In conclusion, my journey at Pickl.AI has been nothing short of extraordinary. The hands-on experience during the internship course equipped me with practical data science skills. The project on developing a booking cancellation prediction model challenged me and led to remarkable outcomes.

Contact Details

LinkedIn: https://www.linkedin.com/in/anmolsinghal07/

GitHub: https://github.com/ayeendroid

Email: anmol.workspace@gmail.com