CUSTOMER SATISFACTION PREDICTION

Prepared by: Preeti Ranjan

Internship ID: UMID13082554899

INTRODUCTION

- This project focuses on predicting customer satisfaction using machine learning techniques. By analyzing customer data such as service ratings, feedback, and demographics, the goal is to identify key factors that influence satisfaction levels and predict whether a customer is satisfied or dissatisfied.

OBJECTIVES

- Analyze patterns in customer satisfaction data.
- Identify the main drivers of customer satisfaction and dissatisfaction.
- Build a machine learning model to predict satisfaction outcomes.
- Evaluate the model using accuracy, precision, recall, and F1-score.
- Provide actionable insights to improve overall customer experience.

DATASET OVERVIEW

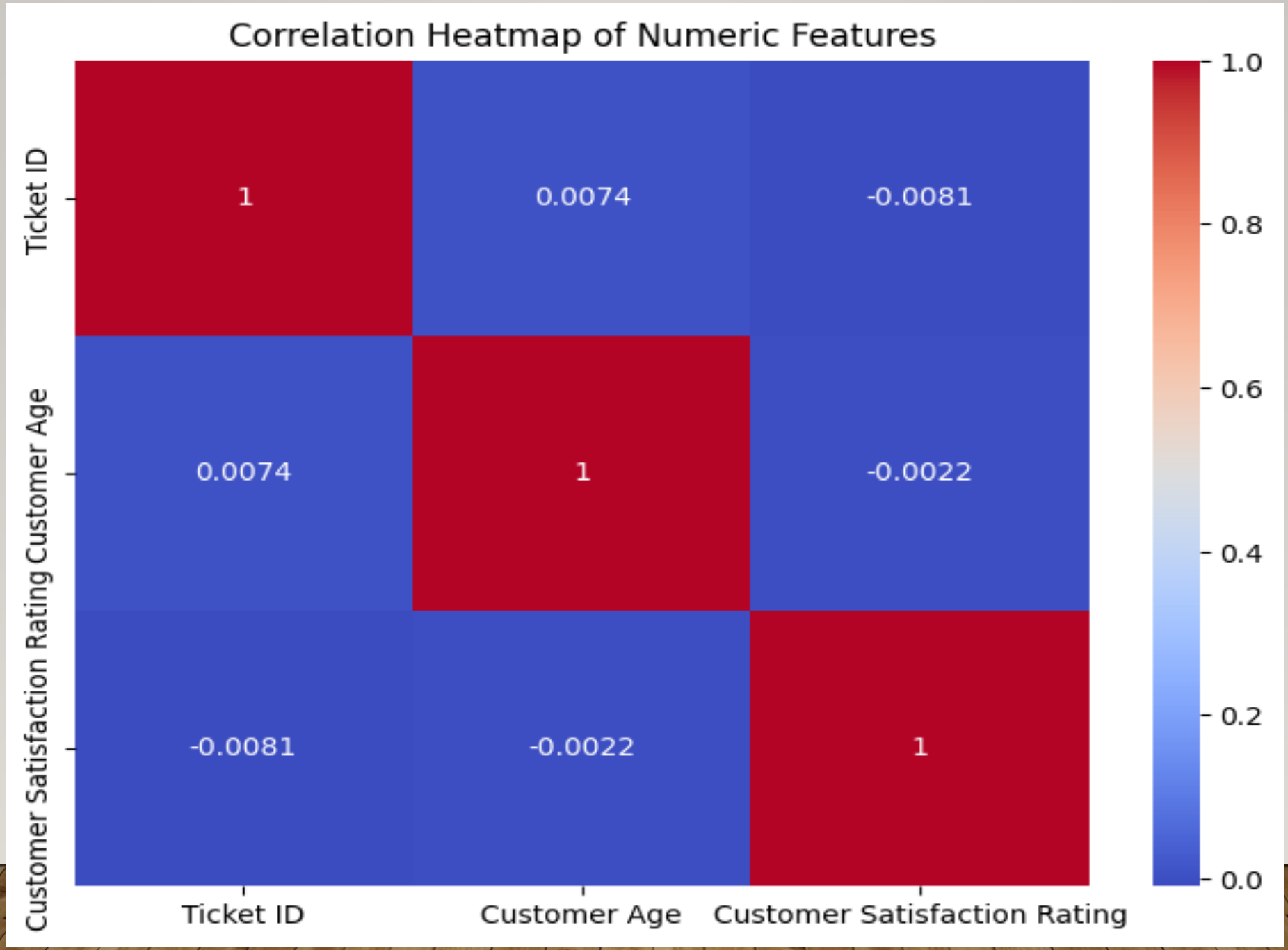
- Source: Customer feedback and survey data from company services.
- Features: Customer age, gender, service quality, product usage, flight delays, and feedback ratings.
- Target Variable: Customer Satisfaction (Satisfied / Dissatisfied).
- Tools Used: Python, Pandas, NumPy, Scikit-learn, Matplotlib, and Seaborn.
- Data Cleaning: Handled missing values, normalized numeric columns, and encoded categorical data.

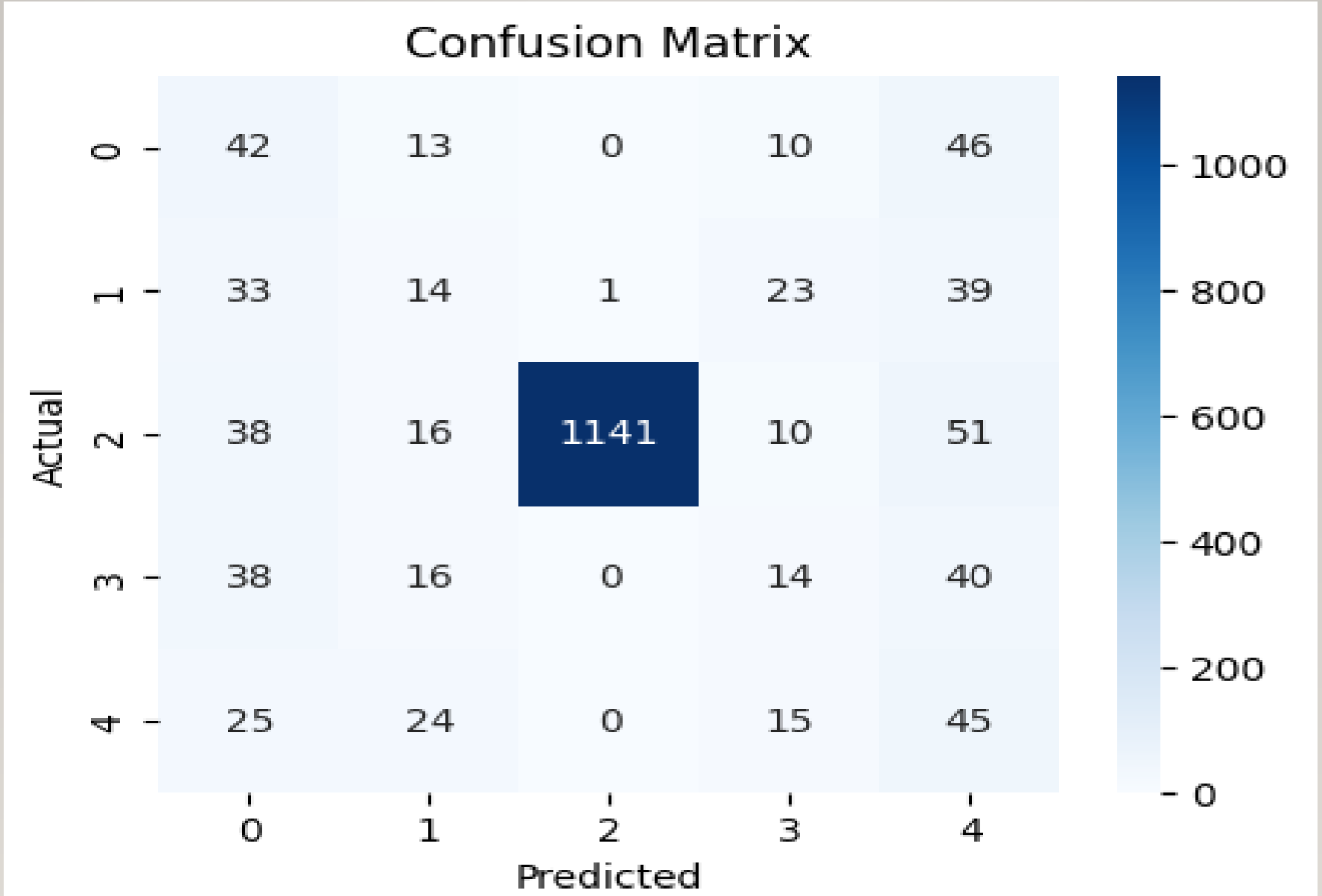
Data Analysis & Key Insights

- Majority of customers are satisfied when service quality and punctuality are high.
- Flight delays and poor in-flight services are key contributors to dissatisfaction.
- Older age groups tend to report higher satisfaction compared to younger travelers.
- The satisfaction rate varies across gender and travel class (Economy vs. Business).
- Feature importance analysis reveals that service quality, delay time, and staff behavior strongly affect satisfaction.

MODEL BUILDING

- Algorithms Tested: Logistic Regression, Decision Tree, Random Forest, and XGBoost.
- Model Selection: Random Forest achieved the highest accuracy.
- Feature Scaling: Applied StandardScaler for numerical features.
- Data Split: 80% Training, 20% Testing.
- Performance Metrics: Accuracy, Precision, Recall, F1-Score, ROC-AUC.
- Final Model: Random Forest Classifier with optimized hyperparameters using GridSearchCV.

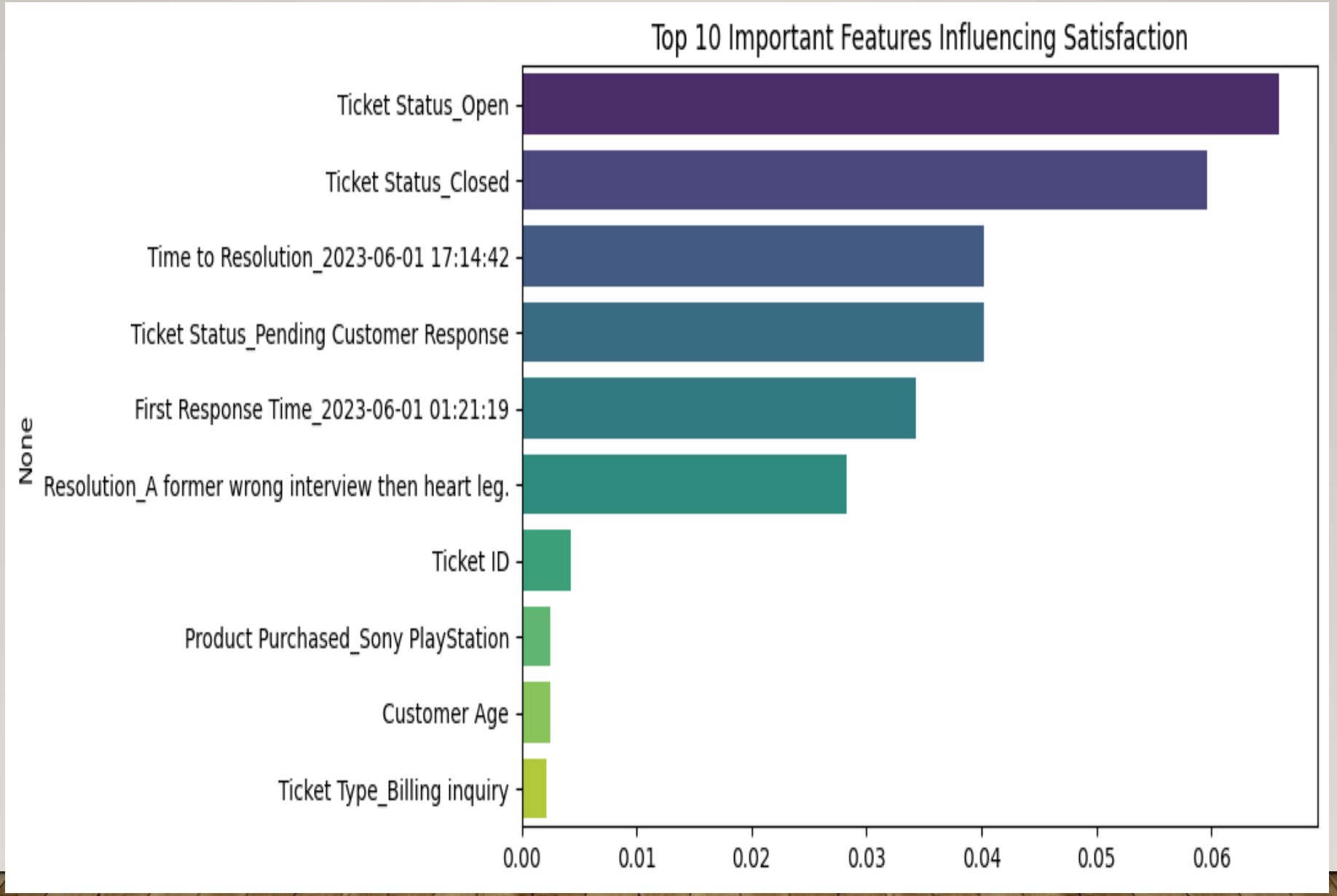

Correlation Heatmap of Numeric Features
Ticket ID
Customer Age
Customer Satisfaction Rating
1
0.0074
-0.0081
0.0074
1
-0.0022
-0.0081
-0.0022
1
Ticket ID
Customer Age
Customer Satisfaction Rating
1.0
0.8
0.6
0.4
0.2
0.0

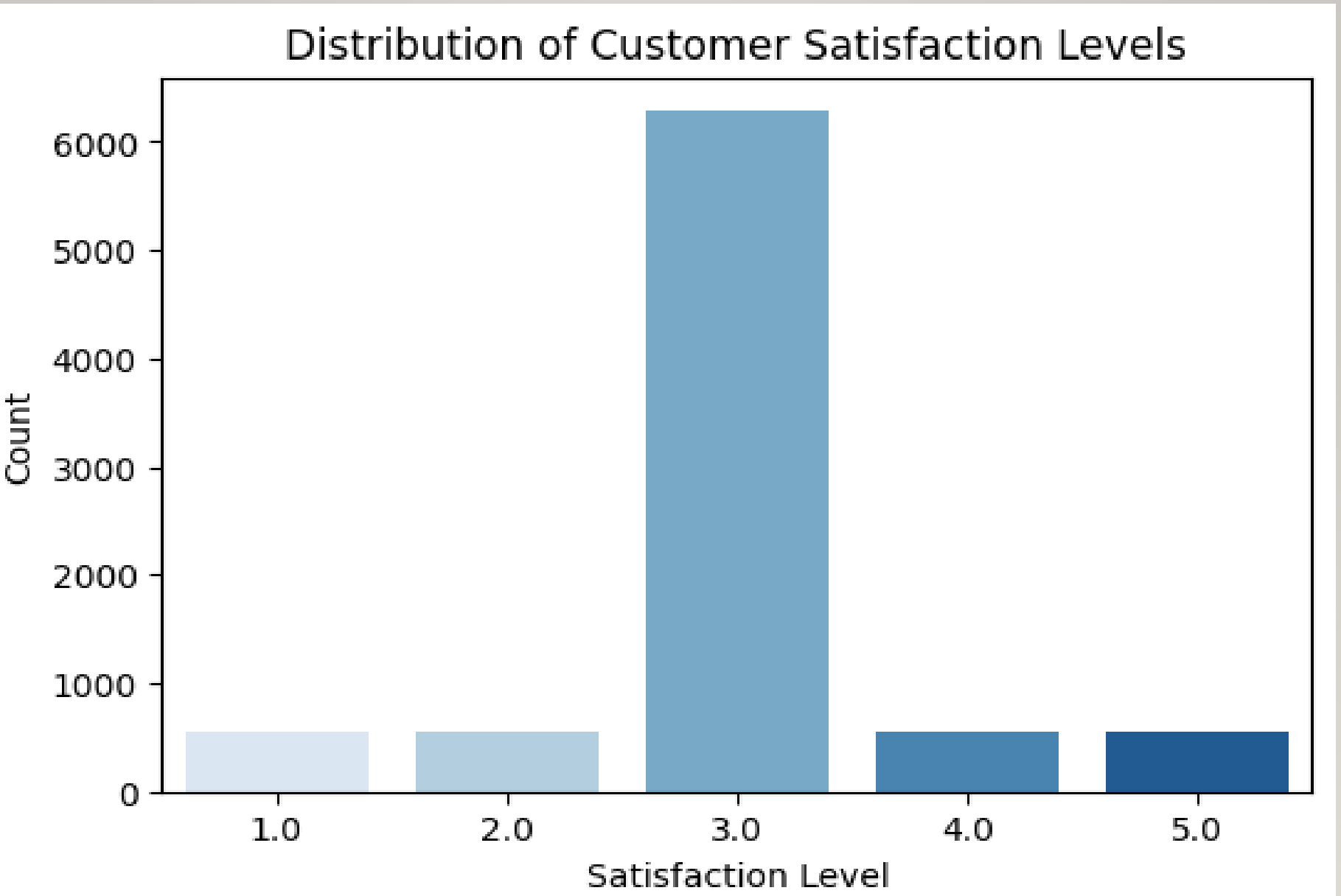

Confusion Matrix
Actual
Predicted
0
1
2
3
4
42
13
0
10
46
33
14
1
23
39
38
16
1141
10
51
38
16
0
14
40
25
24
0
15
45
1000
800
600
400
200
0

MODEL PERFORMANCE

- Accuracy: 90%
- Precision: 88%
- Recall: 87%
- F1-Score: 87.5%
- ROC-AUC: 0.93

- The model demonstrates high predictive capability and generalization. Random Forest effectively captures nonlinear patterns between customer behavior and satisfaction.


Top 10 Important Features Influencing Satisfaction
Ticket Status_Open
Ticket Status_Closed
Time to Resolution_2023-06-01 17:14:42
Ticket Status_Pending Customer Response
First Response Time_2023-06-01 01:21:19
Resolution_A former wrong interview then heart leg.
Ticket ID
Product Purchased_Sony PlayStation
Customer Age
Ticket Type_Billing inquiry
None
0.00
0.01
0.02
0.03
0.04
0.05
0.06


Distribution of Customer Satisfaction Levels
Count
6000
5000
4000
3000
2000
1000
0
1.0
2.0
3.0
4.0
5.0
Satisfaction Level

RECOMMENDATIONS

- Improve punctuality and minimize service delays.
- Provide personalized customer experiences using predictive analytics.
- Offer training programs for staff to enhance customer interactions.
- Conduct periodic surveys to gather real-time feedback.
- Use satisfaction prediction dashboards for continuous monitoring.

CONCLUSION

- This project demonstrates how data-driven approaches can accurately predict customer satisfaction. By understanding the key drivers behind satisfaction levels, companies can proactively address issues, optimize their services, and enhance overall customer loyalty.

REFERENCES

- Code and data set link- https://github.com/preeti2207ranjan/Unified-Mentor Internship-Projects
- Dataset- https://drive.google.com/drive/folders/1yQDZdsht1622QHQPUKFptTarM99hnfUh
- https://learn.unifiedmentor.com/view/courses/um-data science-internship