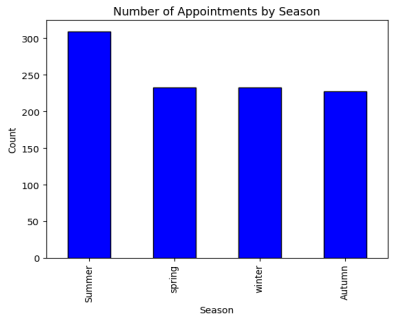
Healthcare No-Show Prediction Report

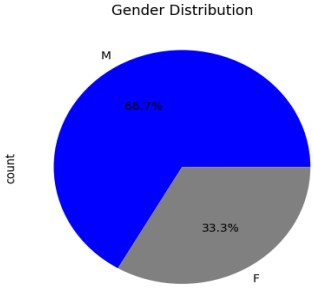
Introduction

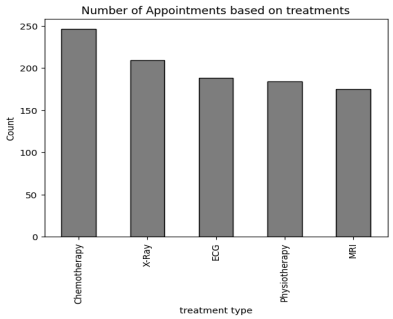
This project analyzes healthcare appointment data to predict whether a patient will show up for their scheduled appointment using machine learning techniques.

Abstract

Missed appointments are costly for healthcare providers. By identifying key predictors such as day of the week, weekend status, and appointment cost, this analysis helps optimize scheduling and resource allocation.





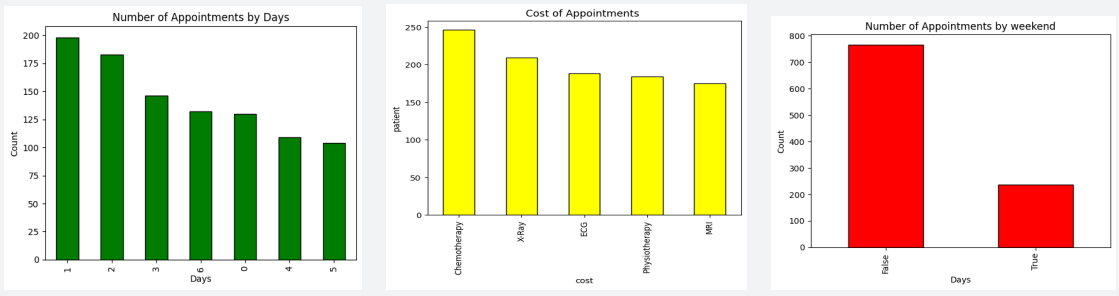


Tools Used

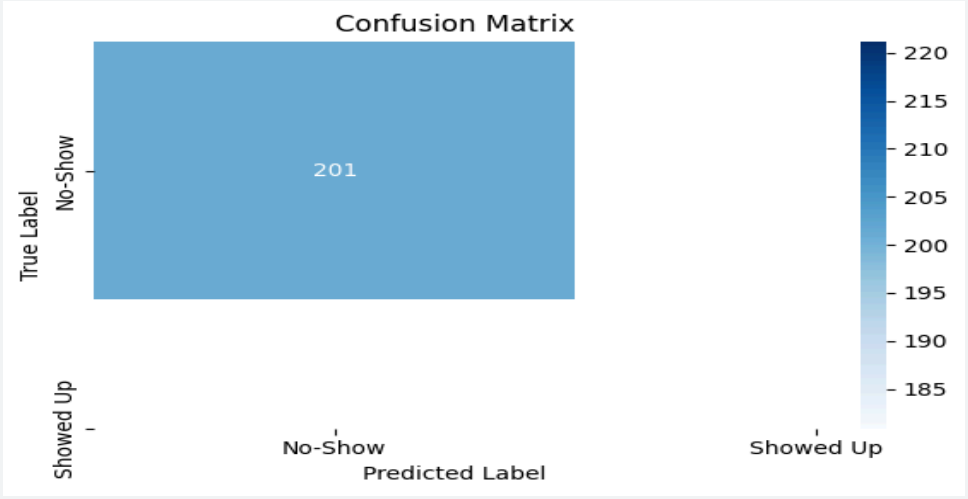
- *Python (Pandas, Scikit-learn, Seaborn, Matplotlib)*
- *Kaggle Notebook*
- *Machine Learning Algorithms (Random Forest)*
- *Power BI for interactive dashboards*

Steps Involved in Building the Project

- ***Data Collection:*** *Merged appointment, patient, doctor, billing, and treatment datasets* ***appointments.info() ,billing.info() ,doctors.info(),patients.info(), treatments.info() to merge_data()***
- ***Data Preprocessing:*** *Cleaned null values, normalized status, encoded categorical features.*
- ***Feature Engineering:*** *Extracted 'days_of_week', 'is_weekend', and treatment costs*



- ***Model Training:*** *Used Random Forest Classifier with train-test split and stratification.*
- ***Evaluation:*** *Evaluated with accuracy, confusion matrix, and classification report*
- ***Visualization:*** *Visualized prediction performance using confusion matrix heatmap*



Conclusion

The trained Random Forest model achieved high accuracy in predicting patient attendance. This solution aids in reducing missed appointments, improving clinic efficiency, and enhancing patient care.