Healthcare Appointment No-Show Prediction

Introduction

Missed healthcare appointments pose significant challenges to both healthcare providers and patients, leading to inefficiencies, wasted resources, and delays in patient care. This project focuses on predicting patient no-shows to optimize scheduling and improve operational efficiency in healthcare facilities. By leveraging historical appointment data, this project identifies patterns and factors contributing to missed appointments.

Abstract

The project utilizes historical healthcare appointment data to predict whether patients are likely to miss their scheduled appointments. Through data preprocessing, exploratory data analysis, and machine learning, a decision tree model was developed to forecast no-shows. Key factors such as age, SMS reminders, and appointment weekdays were analyzed to understand their impact on attendance. The insights are visualized in a Power BI dashboard to aid healthcare providers in effective scheduling and resource allocation.

Tools Used

- Programming Language: Python
- Libraries: Pandas, NumPy, Scikit-learn
- Data Visualization: Power BI
- Other Tools: Excel (for initial data exploration)

Steps Involved in Building the Project

1. Data Collection: Gathered appointment data including patient demographics, appointment details, and reminders.
2. Data Cleaning & Preprocessing: Handled missing values, converted categorical variables, and formatted date and time fields.
3. Exploratory Data Analysis (EDA): Investigated trends in no-shows based on age, gender, appointment weekday, and SMS reminders using descriptive statistics and visualizations.
4. Model Development: Built a decision tree classifier to predict the likelihood of no-shows, optimizing hyperparameters for better accuracy.
5. Evaluation: Assessed model performance using metrics such as accuracy, precision, recall, and F1-score.
6. Visualization & Reporting: Created a Power BI dashboard to visualize key insights, enabling healthcare staff to identify high-risk appointments and optimize scheduling.

Conclusion

The project successfully predicts patient no-shows and highlights key factors influencing appointment attendance. By combining Python-based machine learning models with interactive Power BI dashboards, healthcare providers can proactively manage appointments, reduce resource wastage, and improve patient care. Future enhancements may include integrating real-time notifications and predictive analytics for patient engagement to further reduce missed appointments.