Healthcare Appointment No-Show Prediction - Project Report

1. Introduction

The rising number of missed medical appointments impacts both healthcare efficiency and patient well-being. This project aims to predict appointment no-shows using historical data, enabling better scheduling and patient engagement.

2. Abstract

We analyzed a real-world healthcare dataset to identify patterns behind appointment no-shows. A Decision Tree Classifier was built to predict attendance, and Power BI was used for trend visualization. Insights like age, day of the week, and SMS reminders were key factors. Based on the analysis, we offer scheduling optimization recommendations.

3. Tools Used

- Python: Pandas, Scikit-learn for data processing and model building
- Power BI: For interactive dashboard and trend insights

4. Steps Involved in Building the Project

- Data Import & Cleaning: Loaded appointment dataset, handled missing values, encoded categorical variables, and converted date formats.
- Exploratory Data Analysis (EDA): Identified significant factors affecting no-shows such as SMS reminders, age group, neighborhood, and appointment weekday.
- Model Building: Trained a Decision Tree model to classify whether a patient would show up based on features like age, gender, SMS received, and scholarship.
- Dashboard Creation: Used Power BI to visualize no-show trends by age, weekday, gender, and SMS.
- Optimization Recommendations: Focused on sending reminders to high-risk patients, avoiding overbooking on high no-show days, and prioritizing age groups with higher attendance.

5. Conclusion

The prediction model achieved meaningful accuracy and revealed valuable patterns. This can guide clinics to proactively manage their appointment schedules, improve patient turnout, and allocate resources efficiently. Future improvements may involve ensemble methods or time-based modeling.