Healthcare Appointment No-Show Prediction – Project Report

1. Introduction

Missed healthcare appointments are a major challenge, leading to inefficiency, wasted resources, and delays in patient care. Hospitals face increased costs and lower service quality when patients do not attend scheduled visits. The objective of this project is to predict whether a patient will attend or miss an appointment, and provide insights to optimize scheduling.

2. Dataset Description

The dataset includes patient demographics, scheduling details, and attendance records. Key features are Patient ID, Gender, Age, Neighborhood, Appointment Date, Scheduled Date, SMS reminders, and attendance status. The target variable is No-show (Yes/No).

3. Methodology

Data Preprocessing:

- Removed duplicates and irrelevant fields.

- Handled missing values and formatted date features.

- Encoded categorical variables such as Gender, SMS_received, and No-show.

Model Development:

- Applied a Decision Tree Classifier using Scikit-learn.

- Features: Age, gender, neighborhood, weekday, waiting time, SMS reminders, etc.

- Train-test split: 70-30 ratio.

- Evaluation metrics: Accuracy, precision, recall, F1-score.

Visualization:

- Power BI dashboards developed for insights.

- Key analyses: No-show rate by age group, SMS reminder impact, weekday trends, and neighborhood-level variations.

4. Results & Insights

- Decision Tree achieved ~70–75% accuracy.

- SMS reminders helped reduce no-shows, but not fully effective.

- Younger patients and those with long waiting times were more likely to miss appointments.

- Mondays and Fridays recorded higher no-show rates.

- Certain neighborhoods consistently showed higher no-show probabilities.

5. **Recommendations**

- Strengthen SMS reminder strategies for high-risk groups.

- Reduce long waiting times between scheduling and appointment dates.

- Optimize schedules on high no-show days (Mondays, Fridays).

- Allocate resources efficiently using predictive insights.

6. Conclusion

This project highlights how machine learning can support healthcare operations. By predicting no-shows, hospitals can enhance scheduling efficiency, reduce costs, and improve patient outcomes. The decision tree model combined with Power BI dashboards provides actionable insights for hospital administrators. Future work could explore advanced models like Random Forest or XGBoost, and integrate real-time appointment data.