Healthcare Appointment No-Show Prediction

1. Introduction

Missed medical appointments, or no-shows, pose a significant burden on healthcare systems. They lead to wasted resources, extended patient wait times, and poor health outcomes. This project aims to address this challenge by predicting the likelihood of a patient not attending their scheduled appointment using machine learning and data visualization.

2. Abstract

This project leverages a real-world dataset of over 100,000 medical appointments to identify patterns and predict patient no-shows. A Decision Tree Classifier is used to model no-show behavior, while Power BI is employed to create an interactive dashboard for insights. The goal is to assist healthcare providers in reducing appointment gaps by identifying high-risk patients and taking preventive action such as SMS reminders or scheduling optimizations.

3. Tools Used

- Python Libraries: Pandas, NumPy, Sklearn, Matplotlib, Seaborn
- Machine Learning: Decision Tree Classifier
- Visualization & BI: Power BI
- IDE & Platform: Jupyter Notebook, Power BI Desktop

4. Steps Involved in Building the Project

A. Data Cleaning and Preparation
- Removed duplicates and irrelevant features (e.g., PatientID)
- Converted date columns to datetime format
- Extracted day of the week and calculated waiting time between scheduling and appointment

B. Exploratory Data Analysis (EDA)
- Identified key variables affecting no-shows such as:
  * Age
  * SMS_received
  * Appointment day
  * Time gap between scheduling and appointment

C. Model Building
- Used Decision Tree Classifier for prediction
- Achieved ~59% accuracy
- Generated confusion matrix and classification report

D. Power BI Dashboard
- Created visuals including:
  * KPIs (No-show rate, Total Appointments, Show-up %)
  * Bar and donut charts for age groups, SMS impact, weekdays
  * Matrix and slicers for filtering by gender, neighborhood, scholarship, etc.
- Dashboard enables healthcare decision-makers to identify actionable trends

Healthcare Appointment No-Show Prediction

E. Optimization Insights
- SMS reminders boost attendance
- Friday appointments show high no-show risk
- Long scheduling gaps lead to poor turn-up

5. Conclusion

The project successfully demonstrates how predictive modeling and visual analytics can be used to anticipate patient behavior and optimize healthcare scheduling. By identifying high-risk patients and visualizing influential factors, healthcare providers can reduce no-shows, improve patient care, and increase operational efficiency.