Healthcare Appointment No-Show Prediction

Introduction:
Missed medical appointments, also known as patient no-shows, create operational and financial challenges for healthcare systems. High no-show rates lead to wasted resources, increased waiting times, and reduced access for other patients. This project focuses on predicting no-shows using machine learning and providing insights for scheduling optimization.

Abstract:
The objective of this project is to build a predictive model that determines whether a patient will attend or miss their appointment. We analyzed factors such as SMS reminders, patient age, waiting days, and weekday of appointment. Using Python and Scikit-learn, we trained a decision tree classifier and evaluated its performance. Additionally, a Power BI dashboard was created to visualize trends and provide decision-making support. Optimization strategies are proposed to reduce no-show rates and improve clinic efficiency.

Tools Used:
- Python (Pandas, Scikit-learn, Matplotlib, Seaborn)
- Power BI (interactive dashboards)
- Jupyter/Colab for modeling and analysis

Steps Involved in Building the Project:
1. Imported and cleaned the dataset (removed invalid ages, parsed dates, standardized columns).
2. Feature engineering: waiting days, age bands, weekday of appointment, SMS reminders.
3. Performed exploratory data analysis (EDA) to identify patterns in no-shows.
4. Trained a Decision Tree model and evaluated using confusion matrix, precision, recall, and ROC-AUC.
5. Exported predictions for Power BI dashboard visualization.
6. Built Power BI dashboards with KPIs, trend charts, confusion matrix, and risk distribution.
7. Derived optimization recommendations for reducing no-shows.

Conclusion:
The analysis revealed that SMS reminders, appointment waiting time, and patient age significantly influence no-show rates. The decision tree model provided interpretable predictions that were integrated into a Power BI dashboard for practical insights. Key optimization recommendations include sending additional reminders to high-risk patients, reducing waiting times, and applying overbooking strategies for slots with high predicted no-show rates. This project demonstrates how predictive analytics and business

intelligence can enhance operational efficiency in healthcare scheduling.