### REPORT ON HEALTH CARE APPOINTMENTS PREDICTION

### Introduction

In healthcare, patient no-shows cause wasted time, poor resource utilization, and reduced care efficiency. Hospitals struggle to manage these last-minute absences. To address this, our project aims to analyze past appointment data and build a predictive model that can identify patients who are likely to miss appointments.

### Abstract

This project analyzes a large dataset of patient appointments using Python and machine learning. We processed and cleaned the data, engineered useful features (like waiting days), and trained a classification model to predict appointment attendance. A Power BI dashboard was also created to visualize important patterns, such as the impact of SMS reminders, age, weekday trends, and waiting times on no-shows. The final model achieved ~77% accuracy, offering healthcare providers valuable insights to reduce no-shows and improve operational efficiency.

**Tools Used**

| **Tool / Library** | **Purpose** |
| --- | --- |
| Python (pandas, seaborn) | Data cleaning & analysis |
| scikit-learn | ML model (Decision Tree) |
| Power BI | Visualization & Dashboard |
| Jupyter/IDLE | Code execution |

**Steps Involved**

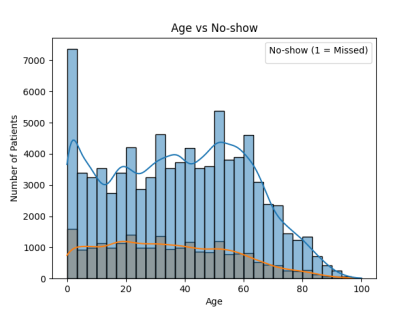
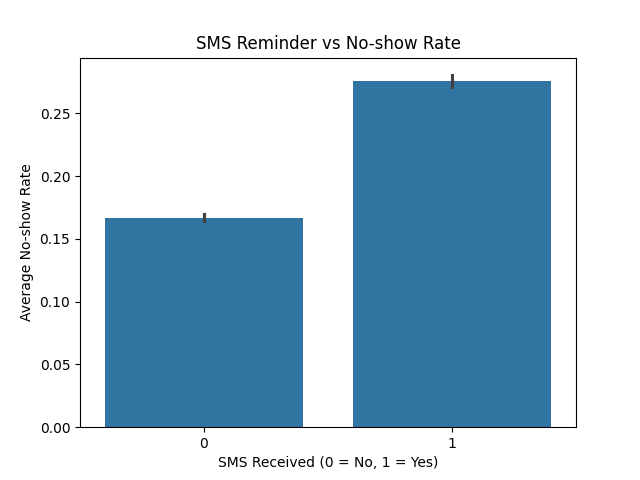
1. **Data Loading & Cleaning**
   * Handled nulls, converted date columns, removed invalid age entries.
2. **Feature Engineering**
   * Created new column: WaitingDays (difference between Scheduled & Appointment day)
   * Encoded No-show column to binary (0 = Show, 1 = No-show)
3. **Exploratory Data Analysis (EDA)**
   * Visualized relationships with age, weekday, SMS, etc. using seaborn
4. **Model Building**
   * Used DecisionTreeClassifier from sklearn
   * Split data (80/20), trained & evaluated the model
   * Accuracy: **77.14%**
5. **Power BI Dashboard**
   * Created visuals: Donut (No-show %), Bar (Age-wise), Column (Weekday), Scatter (Waiting days), and slicers (Gender, Scholarship, SMS)

**Conclusion**

This project demonstrates how machine learning and data visualization can be combined to tackle real-world problems in healthcare. The model helps identify high-risk patients likely to miss appointments, and the dashboard allows hospital staff to explore patterns interactively.

With a few improvements (like patient history, distance, or contact attempts), this system can be implemented to optimize scheduling, improve care delivery, and reduce waste in hospitals

.

**** 

Graph on various data included in CSV file

**PREDICTION ANALYSIS FOR HEALTHCARE APPOINTMENT** 