

# HEALTHCARE APPOINTMENT NO-SHOW PREDICTION

## INTRODUCTION

Missed healthcare appointments, known as “no-shows,” pose a significant challenge for medical professionals, leading to operational inefficiencies and wasted resources. With a growing patient base and tight schedules, optimizing appointment systems has become essential. This project uses historical healthcare appointment data to build a predictive model that identifies patients likely to miss their appointments. Using both machine learning and data visualization tools, the goal is to provide predictive insights and actionable strategies to reduce no-show rates and improve healthcare service delivery.

## ABSTRACT

This project analyzes a dataset of patient appointments to understand behavioral patterns behind no-shows. A decision tree classifier is trained to predict whether a patient will show up, using features like age, gender, health conditions, SMS reminders, and appointment timing. After building the model using Python (Scikit-learn, Pandas, Seaborn), a dynamic dashboard is created in Power BI to communicate findings visually. Insights such as no-show patterns by age group, weekday, neighborhood, and SMS effectiveness are highlighted. Optimization recommendations are derived from trends in patient demographics and scheduling behavior.

## TOOLS USED

- **Python Libraries:**  
Pandas, NumPy – Data cleaning & manipulation  
Matplotlib, Seaborn – Data visualization  
Scikit-learn – Model training and evaluation
- **Power BI:**  
Interactive visual dashboards, data segmentation, KPI cards, pie/donut charts, map analytics, and drill-down filters.
- **Excel:**  
Used as the data source for Power BI import and initial verification.

## STEPS INVOLVED IN BUILDING THE PROJECT

### 1. Data Preprocessing (Python)

- Loaded the Excel dataset (healthappo.xlsx) and cleaned column names.
- Converted date columns to datetime format and extracted Weekday, Month, and DaysGap (days between scheduling and appointment).
- Mapped categorical fields like No\_show, Gender, and SMS\_received to numeric values.
- Created AgeGroup (Child, Adult, Senior) for better visualization.

## 2. Exploratory Data Analysis (Python Visualizations)

- Plotted no-show distribution across gender, age, weekday, and neighborhood.
- Analyzed the impact of SMS reminders on appointment turnout.
- Examined age distribution and feature importance using decision tree output.

## 3. Model Training & Evaluation

- Trained a Decision Tree Classifier using relevant features.
- Achieved performance evaluation through accuracy, classification report, and confusion matrix.
- Identified that DaysGap, Age, and SMS\_received were among the top features influencing no-shows.

## 4. Power BI Dashboard Development

- Created three pages of dashboards:
  - **Page 1:** Demographic trends, weekday analysis, gender-wise comparison, SMS effectiveness.
  - **Page 2:** Neighborhood-wise no-show analysis, income-level correlation, diabetes/hypertension impact.
  - **Page 3:** First-time vs returning patient behavior, interactive filters, drill-down by quarter/month/day.

## 5. Optimization Recommendations

- Send SMS reminders at the optimal lead time (identified from trendline).
- Focus on adult patients in neighborhoods with high no-show rates.
- Prioritize follow-ups for patients with chronic conditions like diabetes or hypertension.
- Introduce weekend appointment slots if feasible, since weekday saturation is evident.

## CONCLUSION

The Healthcare Appointment No-Show Prediction project demonstrates how predictive analytics and interactive dashboards can provide deep insights into patient behaviors. With the help of machine learning, healthcare centers can reduce missed appointments and improve operational efficiency. The dashboard enables administrators to track trends, target interventions, and support data-driven decision-making for better patient outcomes.