# **Healthcare Appointment No-Show Prediction**

**Abstract**

Missed medical appointments, also called no-shows, cause inefficiencies in healthcare delivery, loss of revenue, and disruption in patient care. This project focuses on predicting patient no-shows using machine learning and generating actionable insights with visualization. A Decision Tree and XGBoost-based predictive model was developed using appointment-level data. The analysis considered factors such as age, gender, weekday of appointment, presence of comorbidities, and SMS reminders. A Power BI dashboard was built to provide interactive visual analytics. The outcomes help healthcare providers optimize scheduling and reduce patient no-shows.

**Introduction**

Hospital appointment cancellations or no-shows are a common challenge worldwide. Predictive analytics offers a way to identify patients at high risk of missing their appointments and to take preventive measures, such as reminder SMS or follow-up calls. This project applies machine learning to healthcare appointment data to predict patient behavior. Additionally, trends are visualized through a Power BI dashboard for better decision-making.

**Tools Used**

- Python (Pandas, Scikit-learn, Matplotlib, XGBoost): For preprocessing, model training, and evaluation

.- Power BI: For creating dashboards and visualizing insights.

- Jupyter Notebook: For experimentation and model tuning.

**Steps Involved**

1. Data Import & Cleaning: Loaded dataset (healthcare\_noshows.csv). Removed duplicates, handled missing values, and standardized categorical features.

2.EDA & Feature Engineering: Derived key features such as age group, weekday of appointment, SMS reminders received, and neighborhood. Converted categorical variables into numerical form.

3. Model Development: Trained a Decision Tree Classifier and an XGBoost Classifier. Best performance achieved with tuned threshold (Youden’s J = 0.5861).

4. Model Evaluation: Accuracy: 61.7%, Sensitivity: 57.0%, Specificity: 80.2%, ROC-AUC: 0.686.

5. Dashboard Insights (Power BI): Total appointments: 106,980. Overall No-show rate: 20.26%. Higher no-shows among females and in neighborhoods like Jardim Camburi. SMS reminders reduced no-shows but did not eliminate them. Weekday trends showed higher no-shows on Tuesdays and Mondays.

**Conclusion**

The project demonstrates that machine learning can effectively predict patient appointment no-shows with reasonable accuracy. Insights from the dashboard highlight key demographic and behavioral factors associated with missed appointments.

Recommendations:

- Send targeted reminders to patients predicted as high-risk.

- Schedule flexible time slots in high no-show neighborhoods.

- Optimize staffing on high-risk weekdays.

This predictive approach, combined with interactive dashboards, provides healthcare providers with a practical tool to improve efficiency, reduce cancellations, and enhance patient care.