Healthcare No-Show Appointment Prediction Report

Abstract

Missed medical appointments are a major concern in healthcare, affecting patient outcomes and increasing operational costs. This project predicts whether a patient will show up for their medical appointment using machine learning and visualizes key insights with Power BI.

Introduction

Patient no-shows disrupt healthcare scheduling and efficiency. Predicting appointment attendance can help clinics allocate resources more effectively. In this project, we analyze historical medical appointment data and build a decision tree model to predict patient attendance. Power BI is used to explore visual trends and aid stakeholders with insights.

Tools & Technologies Used

- Python: Data preprocessing, feature engineering, and ML modeling (DecisionTreeClassifier)
- Libraries: pandas, numpy, matplotlib, seaborn, scikit-learn
- Power BI: Interactive dashboards for visualizing patterns
- Jupyter Notebook: Coding and documentation

Steps Involved

1. Data Loading: Dataset from Kaggle loaded and explored.
2. Data Cleaning: Converted date columns, removed invalid ages, and mapped no-show labels.
3. Feature Engineering: Created features like weekday, waiting days; encoded categorical data.
4. Model Training: Trained Decision Tree Classifier with 80/20 train-test split.
5. Evaluation: Used confusion matrix and classification report to evaluate the model.
6. Visualization: Built Power BI dashboards showing age, gender, weekday trends, etc.

Conclusion

The model demonstrated reasonable performance and identified key patterns affecting patient attendance. Power BI visualizations helped identify actionable insights such as high no-show rates on certain weekdays

or among specific age groups. Future improvements can include hyperparameter tuning and integrating SMS reminder behavior as features.