Healthcare Appointment No-Show Prediction – Project Report

Prepared by: Rakghitha S M R

Introduction:

The Healthcare No-show Appointment Prediction project aims to analyze and predict whether a patient will show up for their scheduled medical appointment. No-shows lead to wasted resources and delays in medical care delivery. Using historical appointment data, we aim to build a machine learning model that can accurately predict attendance.

Abstract:

This project utilized a dataset containing appointment and patient details to predict patient behavior. Data preprocessing involved cleaning, encoding categorical values, and extracting features such as weekday from appointment dates. A Decision Tree classifier was used for the prediction task. The final model achieved 73% accuracy, demonstrating potential in helping clinics reduce no-shows and improve scheduling efficiency.

Tools Used:

Python (Pandas, NumPy, Scikit-learn)

Google Colab

Jupyter Notebook

Decision Tree Classifier (Scikit-learn)

Steps Involved in Building the Project:

1. Loaded the dataset in Google Colab.

2. Cleaned the data: Removed unnecessary columns, parsed dates, and encoded categorical values.

3. Feature engineering: Extracted weekday from appointment dates and created dummy variables.

4. Defined features (X) and target (y) variables.

5. Split the dataset into training and testing sets (80/20 split).

6. Trained a Decision Tree classifier model.

7. Evaluated the model using accuracy score and classification report.

Conclusion:

The Decision Tree model was able to predict no-shows with an accuracy of 73%. It showed high reliability in predicting patients who attended their appointments. This project demonstrates the capability of machine learning to address real-world healthcare management challenges and supports the idea of proactive patient engagement.