Introduction

This project focuses on predicting whether a patient will miss a healthcare appointment using machine learning. By analyzing historical data, hospitals can reduce missed appointments and improve scheduling.

Abstraction

Using Python and Power BI, we build a model that predicts patient no-shows based on age, waiting time, SMS reminders, and appointment day. This helps in taking preventive actions like sending reminders or rescheduling.

Tools Used

- **Python**: For data creation, cleaning, feature engineering, and machine learning model building using libraries such as Pandas, NumPy, and Scikit-learn.
- **Power BI**: For interactive visualizations, trend analysis, and dashboard development.
- **Jupyter Notebook / VS Code**: For writing and running Python scripts.
- **CSV File Format**: For data export and transfer between Python and Power BI.

Steps Involved in Building the Project

1. **Data Creation (Dummy Data – 35 Rows)**
    - Generated synthetic healthcare appointment data using Python, simulating real-world scenarios like SMS reminders, age, and appointment days.
2. **Data Preprocessing**
    - Converted date fields, handled types, and created new features like WaitingDays and DayOfWeek.
3. **Exploratory Data Analysis**
    - Examined key columns to understand patterns influencing no-shows.
4. **Model Building**
    - Trained a Decision Tree Classifier to predict the No-show outcome.

- Evaluated model accuracy and added predictions to the dataset.

5. **Data Export**
    - Saved the enriched dataset as a .csv file for use in Power BI.
6. **Power BI Dashboard**
    - Imported the data and created visual insights:
        - SMS vs No-show
        - Age vs Attendance
        - No-show patterns across weekdays
        - Filter slicers for deeper exploration
7. **Recommendation & Optimization**
    - Derived insights to help reduce missed appointments, such as emphasizing SMS reminders or avoiding scheduling elderly patients on high no-show day

Conclusion

The Healthcare Appointment No-Show Prediction project demonstrates how data science and business intelligence can work together to solve real-world problems. By creating a predictive model in Python and delivering visual insights in Power BI, we can support healthcare providers in reducing no-show rates, saving time, and improving overall patient management. Even with a dummy dataset, the approach mimics professional-level healthcare analytics, showing the potential of data-driven decision-making in hospitals.