Healthcare Appointment No-Show Prediction

PROJECT REPORT

Submitted partial fulfilment for the award of the certificate of

Data Analyst Internship

In

Elevate Labs

By

Priyanshu Bagri

OCTOBER 2025

CANDIDATE DECLARATION

It is hereby certified that the work which is being presented in the Project Report entitled **“Healthcare Appointment No-Show Prediction**" is an authentic record of the project carried out during the internship period from September 2025 to October 2025 at Elevate Labs in the capacity of a Data Analyst Intern.

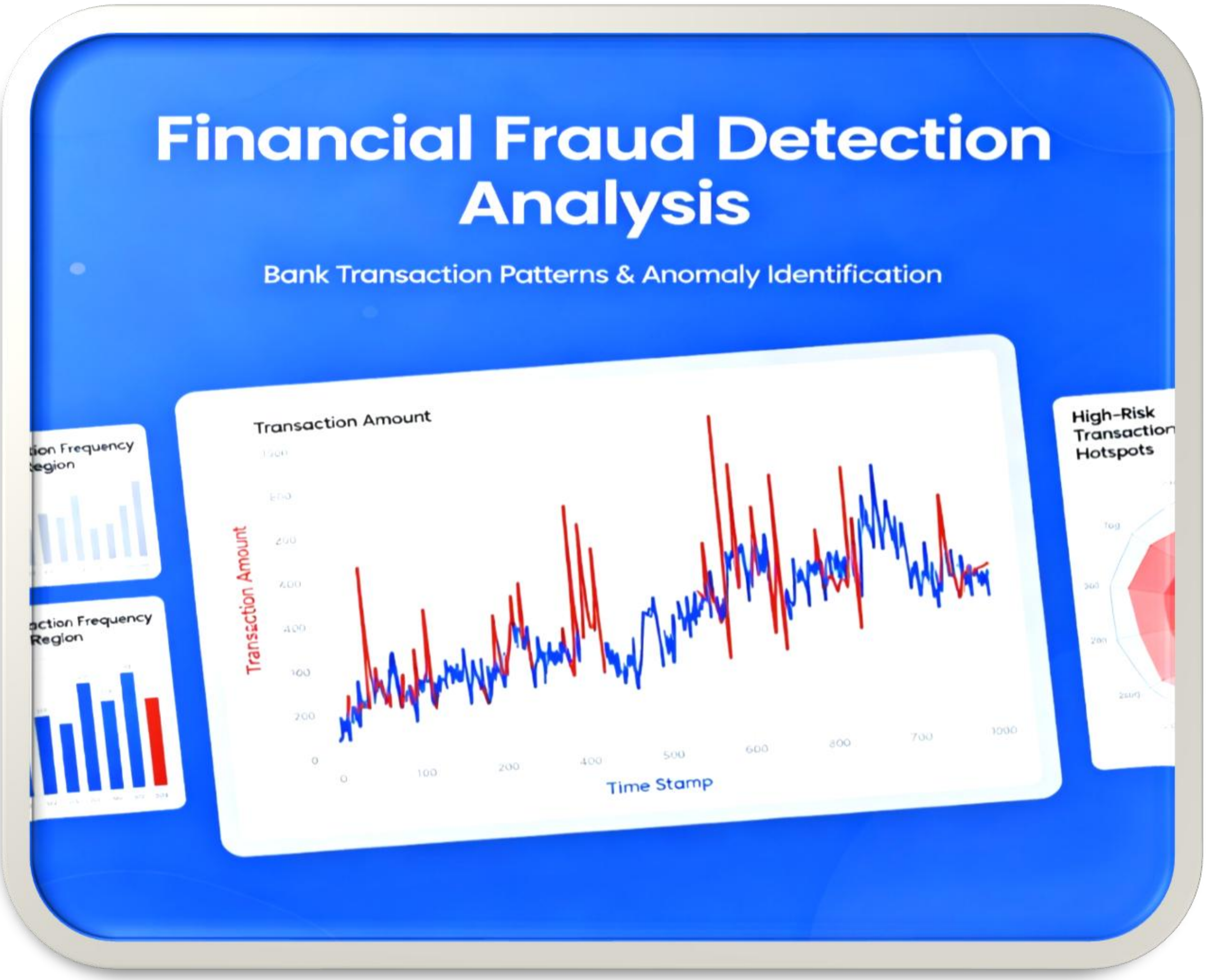
Priyanshu Bagri

TABLE OF CONTENTS

1: INTRODUCTION

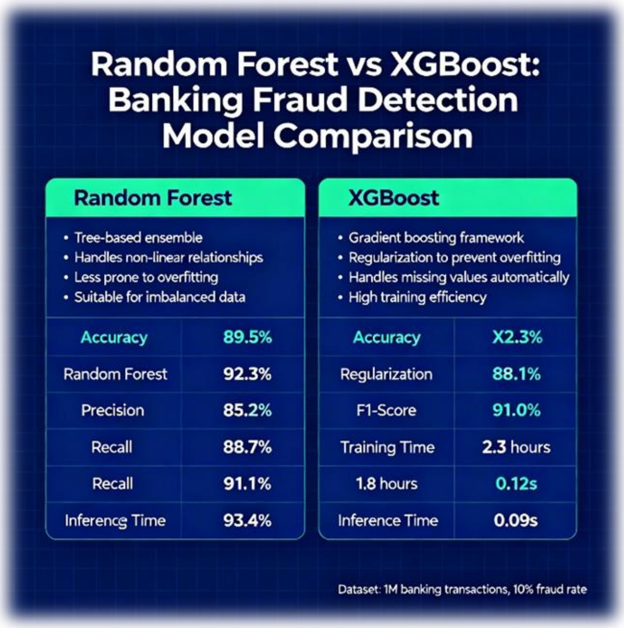
Missed appointments, or “no-shows,” are a core operational challenge in outpatient clinics. They lead to underutilized resources, longer wait times, and reduce the quality of patient care. Predicting no-shows enables clinics to proactively intervene (e.g., reminders, targeted outreach) and optimize scheduling.

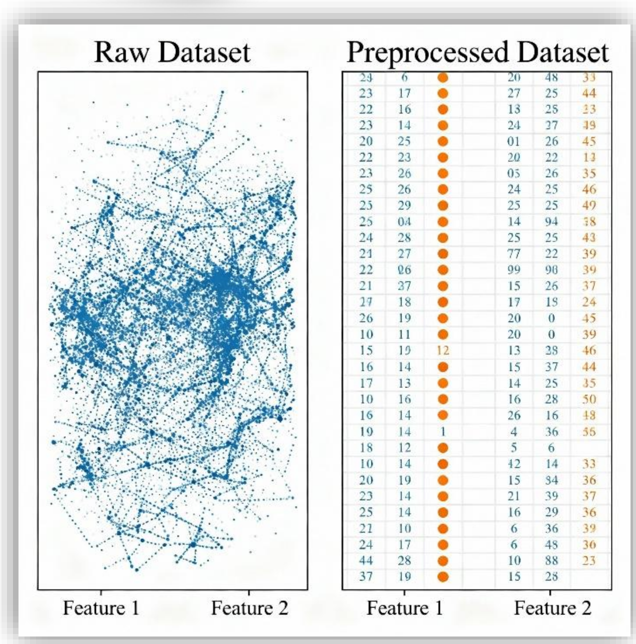
This project leverages Python-based machine learning and Power BI visual analytics to provide a full-cycle, operationally actionable solution for no-show management.



2: Project Objectives

- Predict the likelihood of patient no-shows at the time of scheduling.
- Identify key risk factors associated with missed appointments.
- Enable business users to explore model outputs and operational trends via an interactive dashboard.
- Provide actionable insights for targeted intervention and resource optimization.





3: Dataset Overview

- **Source**: Public no-show dataset (Kaggle, UCI, or clinic data).
- **Features**:
  - **Demographics**: Age (continuous and binned), Gender, Scholarship indicator (low-income).
  - **Medical**: Hypertension, Diabetes, Alcoholism, Handicap.
  - **Appointment Details**: ScheduledDay, AppointmentDay, SMS_received, Neighbourhood, DaysToAppointment, SameDayBooking (engineered).
  - **Target**: No-show (1=missed, 0=attended).
- **Volume**: Over 100,000 appointment records (may vary with cleaning).
- **Imbalance**: Typical no-show rate ~20%.

#	Column	Non-Null Count	Dtype
0	PatientId	110527 non-null	float64
1	AppointmentID	110527 non-null	int64
2	Gender	110527 non-null	object
3	ScheduledDay	110527 non-null	object
4	AppointmentDay	110527 non-null	object
5	Age	110527 non-null	int64
6	Neighbourhood	110527 non-null	object
7	Scholarship	110527 non-null	int64
8	Hipertension	110527 non-null	int64
9	Diabetes	110527 non-null	int64
10	Alcoholism	110527 non-null	int64
11	Handcap	110527 non-null	int64
12	SMS_received	110527 non-null	int64
13	No-show	110527 non-null	object

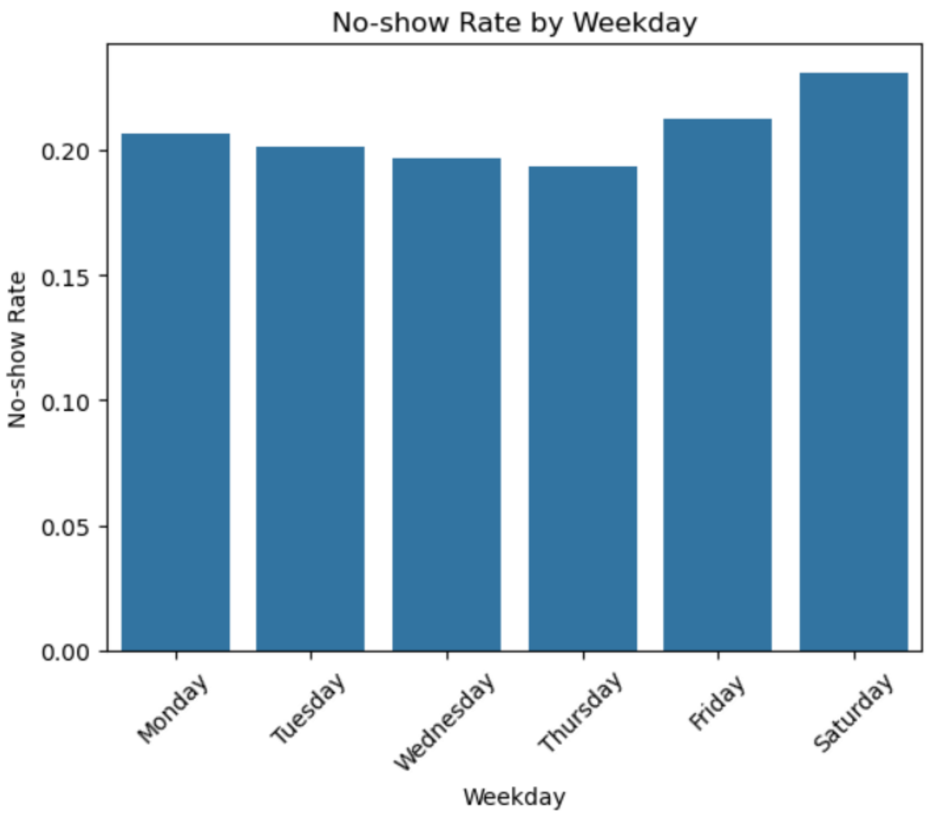
4: Data Cleaning & Preprocessing

4.1 Data Quality Assessment

- Checked for missing and anomalous values.
- Removed duplicate records and appointments with implausible dates or ages.
- Ensured consistent time zone and date formatting.

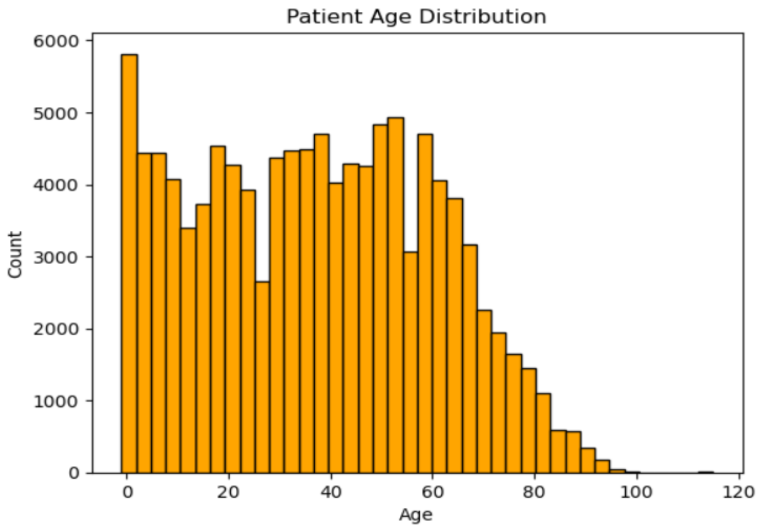
4.2 Feature Engineering

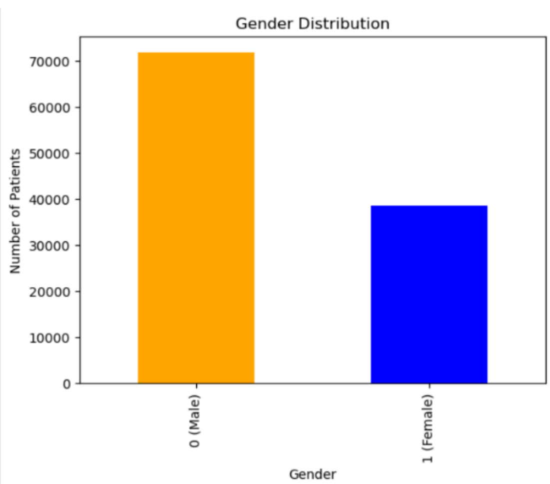
- DaysToAppointment: Calculated as the gap between 'ScheduledDay' and 'AppointmentDay'.
- SameDayBooking: Binary feature for appointments booked and scheduled for the same day.
- Age Groups: Patients divided into strategic age bins using pd.cut.
- Categorical Encoding: Gender and target features mapped to binary (0/1). All engineered features added to the main dataframe for modeling.

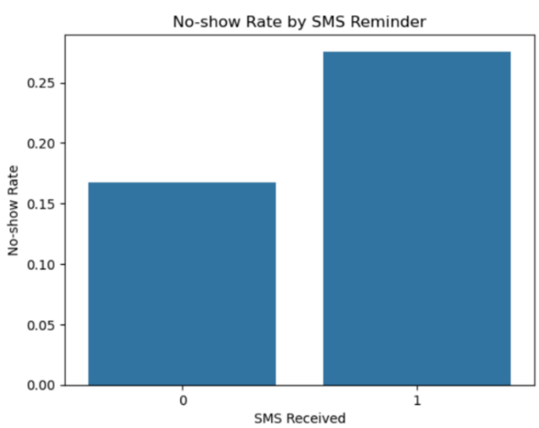


5: Exploratory Data Analysis (EDA)

- **Univariate Analysis:** Age, gender, SMS receipt, medical history distributions.
- **Bivariate Analysis:** No-show rates by age group, gender, days to appointment, weekday, and SMS status.
- **Key Observations:**
    - Younger ages (0–17, 18–30) and same-day bookings show highest no-show risk.
    - SMS reminders are more frequent for at-risk groups, but their isolated effect is limited.
    - Slight gender skew observed in no-show rates, but less significant than age and booking behaviors.
    - No-show rates are cyclic by weekday, peaking around weekends.







6: Modeling Pipeline

6.1 Data Split and Balancing

- Employed a stratified train-test split to ensure outcome proportion consistency.
- Included class weights in model training (where supported) to offset imbalance.

6.2 ML Algorithms Tested

- Baseline: Decision Tree (explainability focus).
- Ensemble: Random Forest and Gradient Boosting (higher stability and predictive power).
  - Used scikit-learn implementations with class_weight='balanced' for robustness.

6.3 Model Evaluation Metrics

- Accuracy: General correctness.
- Precision/Recall (No-show class): Focused on actionable risk flagging.
- F1-Score: Balance between recall (catch more no-shows) and precision (limit false alerts).
- ROC AUC: Probability-based ranking quality across all thresholds.

```
              precision    recall  f1-score   support

           0       0.85      0.71      0.78     17642
           1       0.31      0.50      0.38      4464

    accuracy                           0.67     22106
   macro avg       0.58      0.61      0.58     22106
weighted avg       0.74      0.67      0.70     22106

Confusion Matrix:
 [[12579  5063]
 [ 2216  2248]]
ROC AUC: 0.5898219272678956
```

- **Recall for No-shows is 0.50**: The model captures half of the true no-shows. Precision (0.31) is low, meaning many flagged no-shows are incorrectly predicted, which is common in healthcare no-show prediction due to class imbalance.
- **Overall Accuracy is 67%,** ROC AUC is 0.59: The model is modestly above random guessing, but there’s significant room for improvement with additional features or modeling techniques.
- **Confusion Matrix:** Of all actual no-shows, 2248 were correctly found, but 2216 were missed (false negatives). Of all predicted no-shows, only 2248 out of 5063 were correct.

7: Threshold Tuning & Results

- **Default cutoff** (0.5) led to very low recall (missed most no-shows).
- **Threshold optimization**: Lowered to 0.3, then 0.2, then 0.05—measured trade-offs with confusion matrix and report.
    - At threshold 0.3: Recall for ‘no-show’ improved to 0.48, precision at 0.36 (F1=0.41).
    - Lower thresholds yield higher recall (virtually all no-shows flagged at the cost of increased false positives).

Confusion matrix and metric plots were used to communicate these tradeoffs to clinical stakeholders, allowing for a data-driven decision on “acceptable” risk thresholds.

```
Threshold: 0.5
              precision    recall  f1-score   support

           0       0.80      1.00      0.89     17642
           1       0.50      0.00      0.01      4464

    accuracy                           0.80     22106
   macro avg       0.65      0.50      0.45     22106
weighted avg       0.74      0.80      0.71     22106

[[17623    19]
 [ 4445    19]]
ROC AUC Score: 0.7306153443497291
```

```
Threshold: 0.05
              precision    recall  f1-score   support

           0       0.97      0.32      0.48     17642
           1       0.26      0.96      0.41      4464

    accuracy                           0.45     22106
   macro avg       0.62      0.64      0.45     22106
weighted avg       0.83      0.45      0.47     22106

[[ 5648 11994]
 [  172  4292]]
ROC AUC Score: 0.7306153443497291
```

```
Threshold: 0.3
              precision    recall  f1-score   support

           0       0.86      0.79      0.82     17642
           1       0.36      0.48      0.41      4464

    accuracy                           0.72     22106
   macro avg       0.61      0.63      0.62     22106
weighted avg       0.76      0.72      0.74     22106

[[13887  3755]
 [ 2340  2124]]
ROC AUC Score: 0.7306153443497291
```

(Threshold Results)

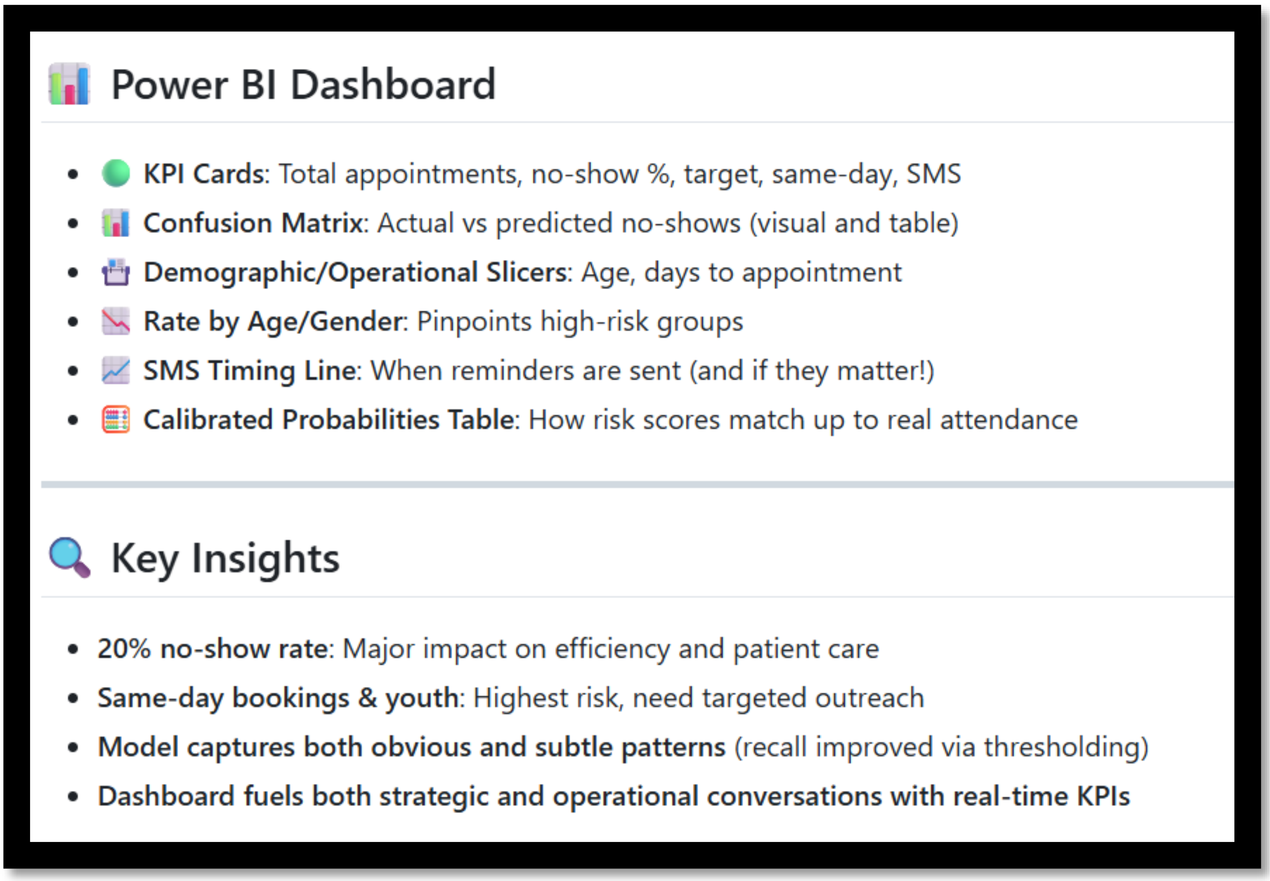
8: Interactive Power BI Dashboard

8.1 Visual Components

- KPI Cards: Total appointments, no-show %, target rates; counts by SMS/same-day booking.
- Confusion Matrix Bar Chart: Displays actual vs predicted, highlighting model's operational performance.
- Demographic Bar/Line Charts: No-show by age, gender, days-to-appointment, and SMS received.
- Probability & Calibration Table: Visualizes model confidence across segments.
- Slicers: Real-time filters for age, days to appointment, and other demographics.

8.2 Dashboard Insights

- Quickly identify segments with highest no-show risk for targeted scheduling/policy changes.
- Track operational KPIs and model trends over time—suitable for recurring clinic management meetings.
- Enable clinic managers or analysts to “slice and dice” no-show risk and see the effect of outreach strategies.



📊 Power BI Dashboard

- 🟢 **KPI Cards**: Total appointments, no-show %, target, same-day, SMS
- 📊 **Confusion Matrix**: Actual vs predicted no-shows (visual and table)
- 🗂️ **Demographic/Operational Slicers**: Age, days to appointment
- 📉 **Rate by Age/Gender**: Pinpoints high-risk groups
- 📈 **SMS Timing Line**: When reminders are sent (and if they matter!)
- 🧮 **Calibrated Probabilities Table**: How risk scores match up to real attendance

🔍 Key Insights

- **20% no-show rate**: Major impact on efficiency and patient care
- **Same-day bookings & youth**: Highest risk, need targeted outreach
- **Model captures both obvious and subtle patterns** (recall improved via thresholding)
- **Dashboard fuels both strategic and operational conversations with real-time KPIs**

9: Key Insights & Recommendations

1. 20% of all appointments result in no-shows.
2. Same-day bookings and youth (ages 0–30) are the highest risk—consider focused interventions and reminders.
3. SMS reminders, while common, are not sufficient alone; behavioral nudges or phone outreach may provide better results for most at-risk groups.
4. Model threshold tuning is essential: lowering cutoff increases recall significantly, albeit at the expense of sending more reminders to patients who would attend.
5. Power BI dashboards empower non-technical users to interpret and act on ML results—improving patient flow and staff resource allocation.

10: Limitations & Next Steps

- **Data Limitations:** No features on prior attendance, external factors (weather, transport, etc.), or clinic-specific workflows.
- **Further Modeling:** Test other algorithms (XGBoost, LightGBM, neural networks), and implement SMOTE or ensemble majority voting.
- **Deployment Potential:** Real-time integration with EMR/EHR systems; design adaptive scheduling and reminder workflows.
- **Evaluation**: Monitor false positive "fatigue" and periodically retrain the model as patient trends and clinic operations evolve.

11: Conclusion

This project demonstrates a full-cycle, real-world application of machine learning for healthcare operations. By combining structured, explainable Python ML, targeted feature engineering, operational threshold tuning, and powerful BI reporting, the project delivers measurable gains for patient management, resource optimization, and proactive healthcare delivery.

Future extensions could include deeper time-series analysis, NLP-driven communications, and automated rescheduling for persistent no-shows

12: References

- Kaggle No-Show Appointments Dataset
- scikit-learn Documentation
- Power BI Official Documentation
- Cited ML research and professional blogs (**DigitalOcean**, [Kaggle Notebooks])