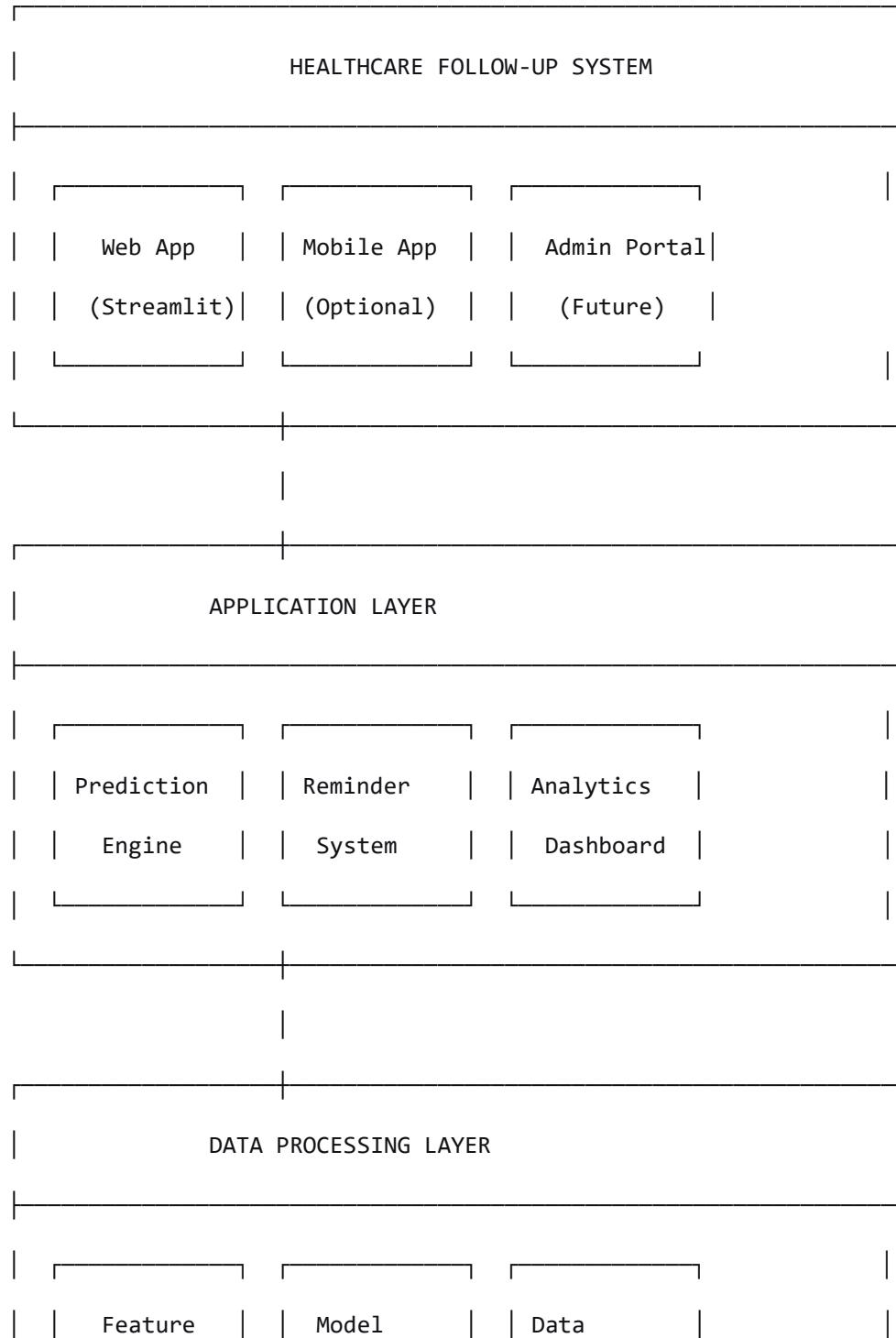
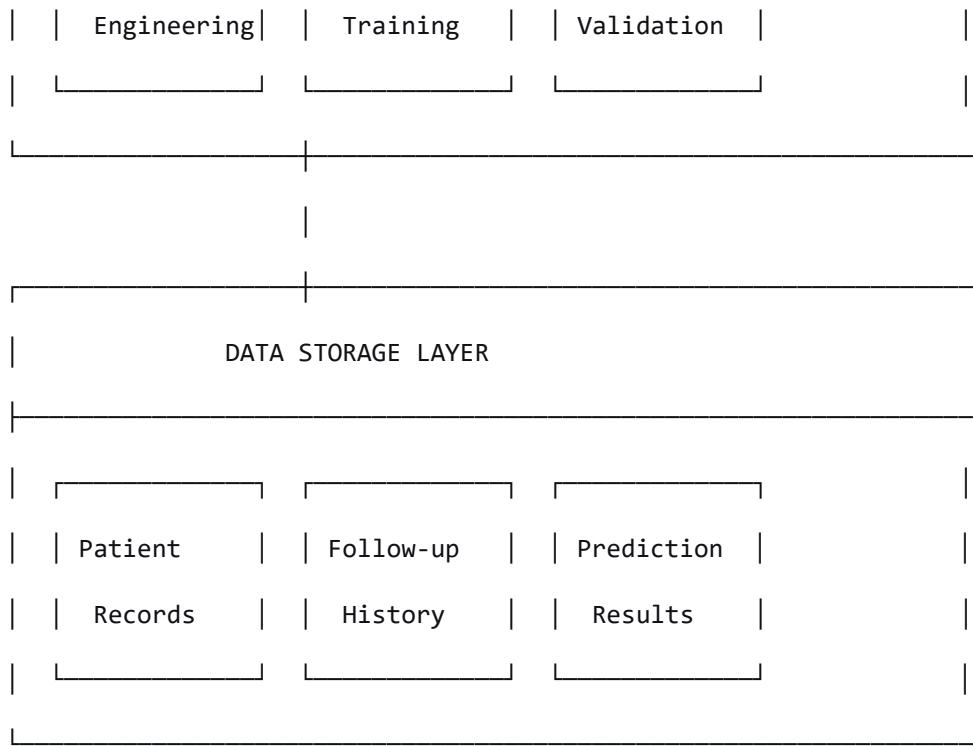


Healthcare Follow-up Default Prediction System - Solution Architecture

1. System Overview



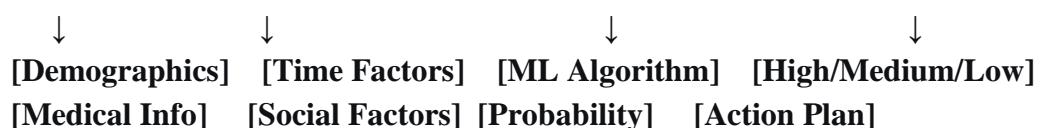


2. Core AI Methods and Technologies

Machine Learning Approach

PREDICTION WORKFLOW:

Patient Data → Feature Extraction → Model Prediction → Risk Assessment



Key AI Methods:

1. Supervised Machine Learning

- Random Forest Classifier
- Logistic Regression
- Gradient Boosting (XGBoost)

2. Feature Engineering

- Time-based features (days overdue, appointment gaps)
- Demographic risk scoring
- Social determinant factors

3. Rule-based Fallback System

- Heuristic scoring for model backup
- Business logic implementation

Technology Stack

FRONTEND: Streamlit (Python web framework)

BACKEND: Python (Scikit-learn, Pandas, NumPy)

DATA: Excel/CSV files (can upgrade to database)

CLOUD: deployment (Streamlit Cloud)

3. Data Requirements and Prerequisites

Required Datasets

PATIENT DEMOGRAPHICS:

- └─ Age, Gender, Occupation
- └─ Insurance status
- └─ Contact information
- └─ Location data (if available)

MEDICAL INFORMATION:

- └─ Diagnosis and condition type
- └─ Medications prescribed
- └─ Treatment complexity
- └─ Follow-up reason

APPOINTMENT DATA:

- └─ Discharge dates
- └─ Follow-up schedule
- └─ Historical attendance
- └─ Cancellation patterns

SOCIAL DETERMINANTS:

- └─ Employment status
- └─ Transportation access
- └─ Social support
- └─ Economic factors

Data Prerequisites

MINIMUM DATA REQUIREMENTS:

- ✓ Patient age and gender
- ✓ Follow-up date and status
- ✓ Basic medical information
- ✓ Contact details

OPTIONAL ENHANCEMENTS:

- ✓ Historical attendance records
- ✓ Socio-economic indicators
- ✓ Geographic data
- ✓ Communication preferences

4. Open Source Resources

Core Libraries

```
python
# Machine Learning
scikit-learn    # Model training and prediction
pandas         # Data manipulation
numpy          # Numerical computations
joblib         # Model serialization

# Web Application
streamlit      # Web interface
plotly         # Interactive visualizations
matplotlib     # Basic plotting
seaborn        # Statistical visualizations

# Data Processing
datetime       # Date/time handling
re             # Regular expressions
os             # File system operations
```

5. System Components Detailed

Prediction Engine Architecture

INPUT FEATURES:

Time Factors	Social Factors	Medical Factors
<ul style="list-style-type: none">• Days overdue• Appointment gap• Seasonality	<ul style="list-style-type: none">• Employment• Insurance• Age group• Contact info	<ul style="list-style-type: none">• Condition• Medications• Complexity• History

PREDICTION PROCESS:

1. Data Validation → 2. Feature Calculation → 3. Model Inference
4. Risk Scoring → 5. Recommendation Generation → 6. Output

Risk Calculation Logic

HIGH RISK PATIENTS (Probability > 60%):

- ✓ Overdue appointments
- ✓ No insurance + unemployed
- ✓ No telephone contact
- ✓ Young age (≤ 25) or elderly (≥ 70)
- ✓ Multiple risk factors combined

MEDIUM RISK (35-60%):

- ✓ Due soon appointments
- ✓ Some risk factors present
- ✓ Partial contact information
- ✓ Moderate social challenges

LOW RISK (<35%):

- ✓ Upcoming appointments
- ✓ Good contact information
- ✓ Stable employment/insurance
- ✓ Reliable history

6. Implementation Roadmap

Phase 1: Basic System (Current)

- ✓ Streamlit web interface
- ✓ Excel-based data storage
- ✓ Heuristic risk prediction
- ✓ Basic visualization
- ✓ Manual data entry

Phase 2: Enhanced Features

- Database integration (SQLite/PostgreSQL)
- Automated data imports
- SMS/email reminders
- Advanced analytics
- Model retraining pipeline

Phase 3: Advanced AI

- Integration with EHR systems
- Natural language processing
- Predictive analytics dashboard
- Mobile application
- API services

7. Data Flow Diagram

1. DATA COLLECTION

↓

2. FEATURE EXTRACTION

- |—— Time-based features
- |—— Demographic features
- |—— Medical features
- |—— Social features

↓

3. MODEL PREDICTION

- └─ Primary: Machine Learning model
- └─ Fallback: Rule-based heuristic

↓

4. RISK CLASSIFICATION

- └─ High risk: Immediate action
- └─ Medium risk: Proactive follow-up
- └─ Low risk: Standard process

↓

5. ACTION RECOMMENDATIONS

- └─ Communication plans
- └─ Escalation procedures
- └─ Resource allocation

8. Success Metrics

PREDICTION ACCURACY:

- Default prediction precision
- Risk classification accuracy
- Model performance metrics

OPERATIONAL IMPACT:

- Reduced missed appointments
- Improved patient outcomes
- Resource optimization
- Staff efficiency gains

PATIENT SATISFACTION:

- Better follow-up adherence
- Improved communication
- Enhanced care continuity