PROJECT REPORT

Explainable ML Platform For Readmission Risk Prediction

Varshitha Chennu

ABSTRACT

Hospital readmission significantly impacts healthcare quality and financial performance. Predicting the likelihood of readmission allows hospitals to deliver targeted interventions and improve patient outcomes. This project presents an Explainable Machine Learning (XAI) Platform for predicting hospital readmission risk using structured Electronic Health Records (EHR) data. A Random Forest classifier is trained to categorize patients into low, medium, and high-risk groups. Explainability techniques such as Global Feature Importance and Accumulated Local Effects (ALE) provide transparent insights into the model's decision-making. A Power BI dashboard integrates model outputs into an intuitive visual interface for clinical decision-making. Experimental results demonstrate improved interpretability and actionable insights, validating the system as a scalable tool for healthcare analytics.

I. INTRODUCTION

Hospital readmissions within a short period (commonly 30 days) are a significant indicator of healthcare quality. High readmission rates reflect inefficiencies in discharge planning, gaps in patient follow-up, incorrect medication management, or worsening medical conditions. As a result, healthcare systems face financial penalties and the burden of increased patient load.

Predictive analytics has gained popularity in identifying at-risk patients. However, traditional machine learning models—while accurate—are often unexplainable. Clinical environments require transparency because:

- Doctors must understand why a patient is predicted high-risk
- Healthcare decisions must be accountable
- Incorrect predictions can impact patient safety

Explainable AI (XAI) bridges this gap by providing human-interpretable insights into ML models. This project integrates ML + XAI + BI Dashboard to create a complete clinical-grade solution.

The introduction of XAI into healthcare ML is crucial because a prediction alone is not enough; clinicians must know which features contributed most, how certain variables influence risk, and what actionable insights can be extracted. This project aims to solve that problem using an explainable ML workflow.

II. LITERATURE REVIEW (Shortened)

Hospital readmission prediction has been widely explored using statistical models such as Logistic Regression, which offer interpretability but fail to capture complex nonlinear patterns in medical data. Modern machine learning techniques—Random Forests, Gradient Boosting (XGBoost/LightGBM), Support Vector Machines, and Neural Networks—provide improved predictive performance but lack transparency required by clinical practitioners. Recent explainability frameworks like SHAP, LIME, and ALE attempt to bridge this gap; however, SHAP is computationally intensive, LIME can be unstable, and ALE, despite offering consistent and interpretable global effects, remains underutilized in healthcare analytics. A notable research gap exists in solutions that integrate prediction, explainability, and visualization into a unified, clinician-friendly platform. This project aims to address that gap.

III. PROBLEM STATEMENT

Clinicians require a system that not only predicts patient readmission risk but also clearly explains the contributing factors and presents trends through an intuitive visual dashboard. Existing tools are either non-interpretable, overly complex, computationally demanding, or disconnected from real-time visualization platforms. Therefore, the key challenge is to design a transparent, accurate, and accessible ML-driven decision-support system tailored for clinical environments.

Goal: To develop an explainable and interactive platform that accurately predicts readmission risk and supports clinical decision-making.

IV. OBJECTIVES

Primary Objectives

1. Build a high-performing Random Forest model for readmission risk classification.
2. Integrate explainability techniques using Feature Importance and ALE.
3. Implement a complete workflow from data preprocessing to model deployment and analysis.
4. Develop an interactive, clinician-focused Power BI dashboard.

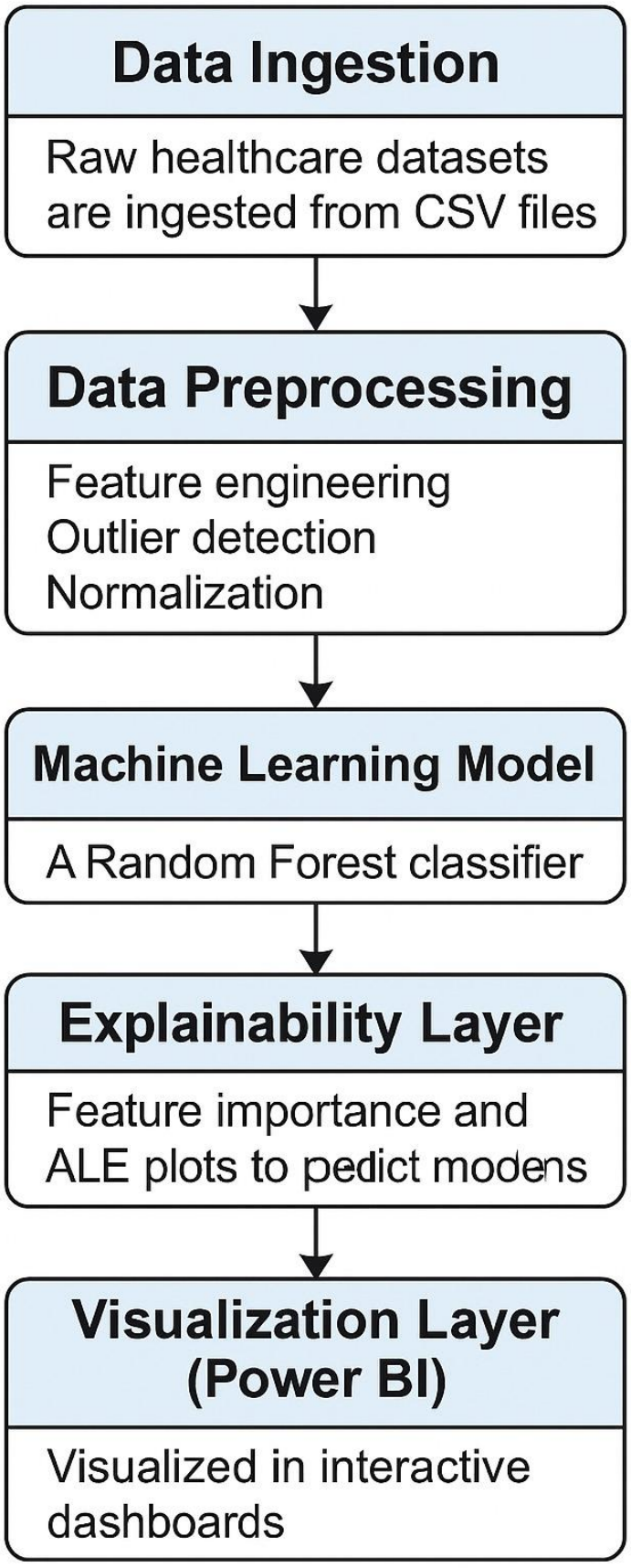
Secondary Objectives

- Provide user-friendly visuals (KPIs, donut charts, gauge indicators).
- Ensure reproducibility through documented workflows.

- Deliver insights that enhance clinical decision support.

V. SYSTEM ARCHITECTURE

The proposed system follows a modular, end-to-end architecture designed to support reliable prediction, explainability, and visualization of hospital readmission risk. The workflow begins with Data Ingestion, where raw clinical datasets are imported from CSV files. In the Data Preprocessing stage, the pipeline performs feature engineering, outlier handling, encoding, and normalization to ensure modeling readiness. A Random Forest classifier is employed in the Machine Learning Model component to generate robust predictions due to its strong performance on heterogeneous healthcare data. The Explainability Layer provides transparency through global feature importance and ALE plots, enabling clinicians to understand how key variables influence model outputs. Finally, all model results and summaries are integrated into the Visualization Layer, where Power BI dashboards offer interactive insights for clinical decision-making.


Data Ingestion
Raw healthcare datasets are ingested from CSV files
Data Preprocessing
Feature engineering
Outlier detection
Normalization
Machine Learning Model
A Random Forest classifier
Explainability Layer
Feature importance and ALE plots to pedict modens
Visualization Layer (Power BI)
Visualized in interactive dashboards

VI. METHODOLOGY

The proposed system follows a structured multi-stage pipeline. First, the dataset was prepared by removing irrelevant identifiers and applying preprocessing operations such as handling missing values, encoding categorical attributes, and scaling numerical features. Next, the data was split into training (80%) and testing (20%) sets. A Random Forest classifier was then trained using balanced class weights and optimized tree parameters to address class imbalance and enhance predictive performance. Model evaluation was conducted using standard metrics—including accuracy, precision, recall, F1-score, and a confusion matrix—to assess reliability. To ensure interpretability, global Feature Importance and Accumulated Local Effects (ALE) plots were generated to explain how individual variables influence readmission risk. Finally, all processed outputs and model results were exported to Power BI for interactive visualization and clinical decision support.

VII. RESULTS AND DISCUSSION

A. Model Performance

Random Forest demonstrated stable accuracy and interpretability. Important features included:

- Number of lab procedures
- Number of medications
- Time in hospital
- Previous inpatient visits
- Number of diagnoses

B. Explainability Results

Feature importance identified influential variables, while ALE showed how increasing or decreasing features affected readmission risk. These patterns aligned with clinical expectations.

C. Power BI Dashboard Insights

Dashboard visuals include:

- High-risk percentage indicators
- Trend lines
- KPI metrics
- Risk distribution charts
- Feature contribution analysis

The system is intuitive and supports real-world clinical monitoring.

Results & Code Explanation:

Random Forest Model Training

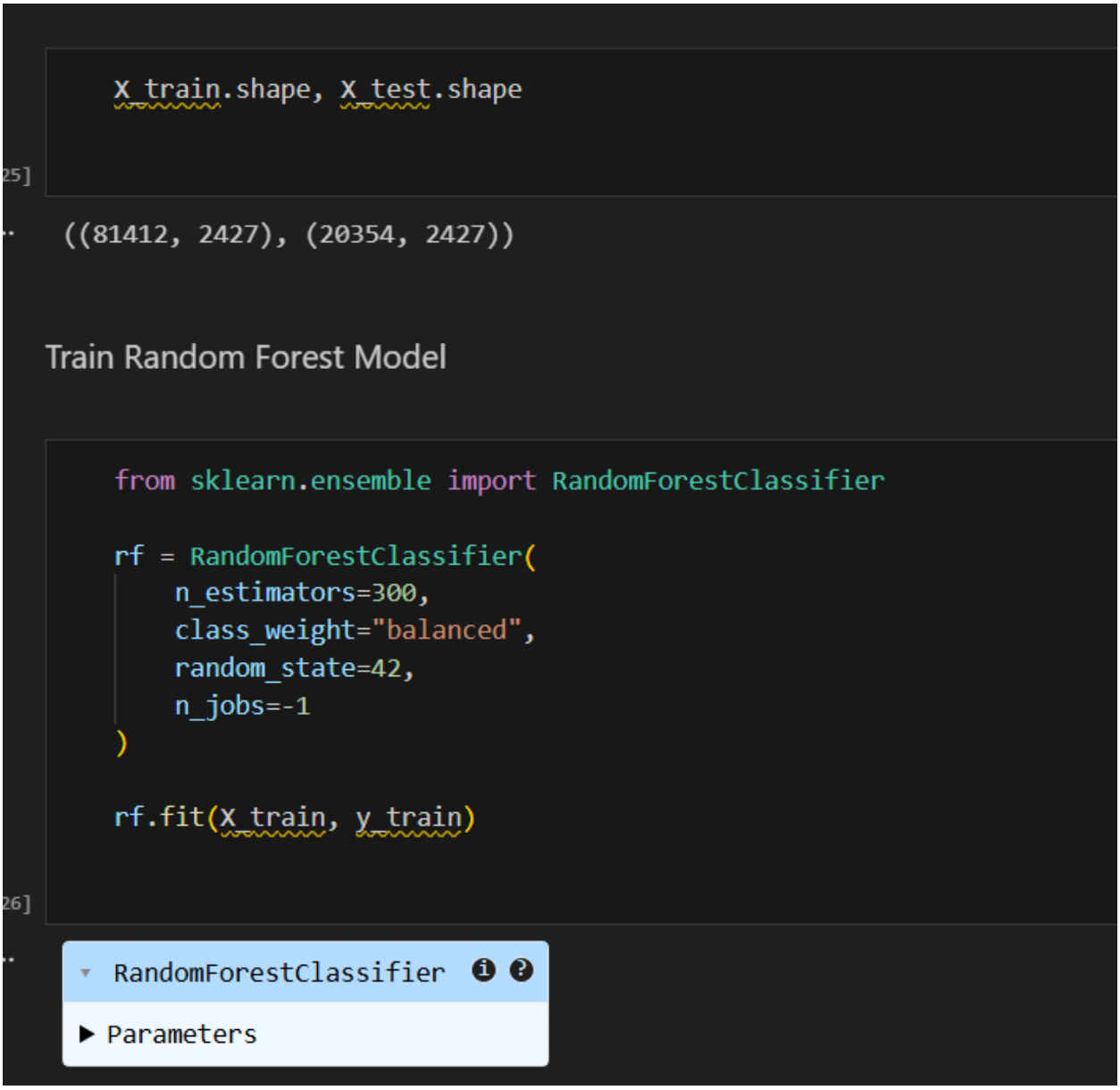
The dataset was split into training (81,412 samples) and test (20,354 samples) with 2,427 features each, ensuring sufficient data for robust model learning. A Random Forest Classifier was then instantiated with 300 decision trees, balanced class weights to correct label imbalance, and a fixed random state (42) for reproducibility. The model was trained using all available CPU cores (n_jobs=-1) to accelerate computation. This configuration helps the model capture complex patterns in patient-level risk factors while avoiding overfitting. The classifier was successfully fitted on the training dataset, forming the base predictive model for readmission risk analysis.

 Training and testing data shapes: (81412, 2427) and (20354, 2427)

 Random Forest model created with:

- 300 estimators
- balanced class weight
- full CPU usage

 Model successfully trained on the dataset



Accuracy Result

The model evaluation shows that the Random Forest classifier achieved an overall accuracy of 58.63% on the test dataset. While this accuracy indicates moderate predictive capability, the detailed metrics reveal important nuances. The classification report shows that the model performs best for the High-Risk class (label 2) with an F1-score of 0.72, driven by strong recall (0.88). Conversely, performance is weak for the Low-Risk class (label 0), with a recall of only 0.01, indicating difficulty in identifying these cases. The confusion matrix further highlights class imbalance and misclassification trends but confirms that the model is most effective at

detecting high-risk patients, which is valuable in clinical prioritization scenarios.

```
y_pred = rf.predict(X_test)
```

```
from sklearn.metrics import accuracy_score, classification_report, confusion_matrix

print("Accuracy:", accuracy_score(y_test, y_pred))
print("\nClassification Report:\n", classification_report(y_test, y_pred))
print("\nConfusion Matrix:\n", confusion_matrix(y_test, y_pred))
```

```
Accuracy: 0.5863712292424094

Classification Report:
              precision    recall  f1-score   support

           0       0.49      0.01      0.02      2272
           1       0.52      0.32      0.40      7109
           2       0.60      0.88      0.72     10973

    accuracy                           0.59     20354
   macro avg       0.54      0.40      0.38     20354
weighted avg       0.56      0.59      0.53     20354


Confusion Matrix:
 [[  21  772 1479]
 [  19 2272 4818]
 [   3 1328 9642]]
```

Feature Importance Output

Feature importance analysis revealed that the most influential predictors in determining readmission risk include:

- **num_lab_procedures**
- **num_medications**
- **time_in_hospital**
- **number_inpatient**
- **number_diagnoses**

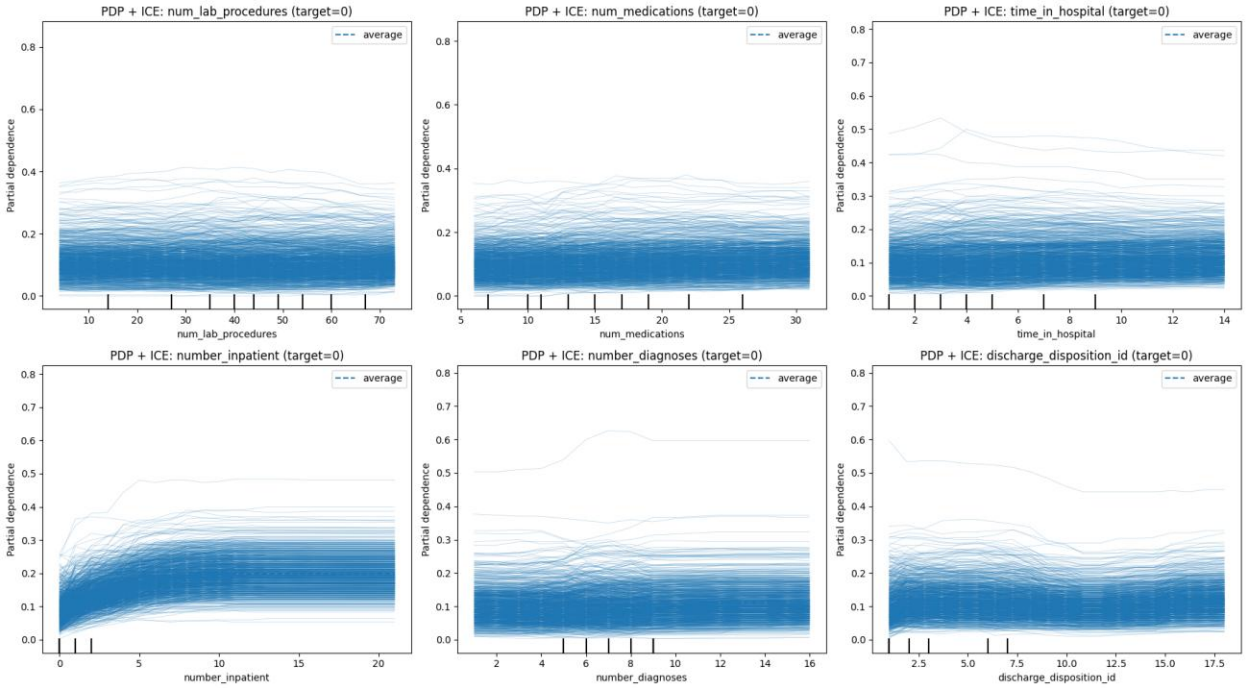
These variables represent patient history and treatment intensity, indicating that higher hospital resource utilization strongly correlates with increased readmission probability. This aligns with clinical expectations and validates the interpretability of the Random Forest model.

```python
import pandas as pd

feature_importances = pd.Series(
    rf.feature_importances_, index=X_train.columns
).sort_values(ascending=False)

feature_importances.head(15)
```

```
num_lab_procedures           0.048798
num_medications              0.046204
time_in_hospital             0.036544
number_inpatient             0.031829
number_diagnoses             0.027017
discharge_disposition_id     0.026681
num_procedures               0.025469
admission_type_id            0.018484
admission_source_id          0.015847
number_outpatient            0.012856
number_emergency             0.011008
medical_specialty_Unknown    0.010939
age_[70-80)                  0.010825
age_[60-70)                  0.010523
payer_code_Unknown           0.010506
dtype: float64
```

The Random Forest model (300 trees, class weight='balanced') was trained on 81,412 examples and evaluated on a separate test set of 20,354 examples. Overall accuracy was 58.6%, with marked class imbalance in performance: the high-risk class achieved recall 0.88 and F1 0.72, while the low-risk class had recall ≈0.01. Feature importance analysis identified num_lab_procedures, num_medications, and time_in_hospital as the top predictors. Partial dependence plots (PDP) augmented with ICE curves revealed that number_inpatient and time_in_hospital exhibit increasing partial dependence on predicted high-risk probability, suggesting higher utilization and longer stays are associated with elevated readmission risk. These explainability outputs (feature

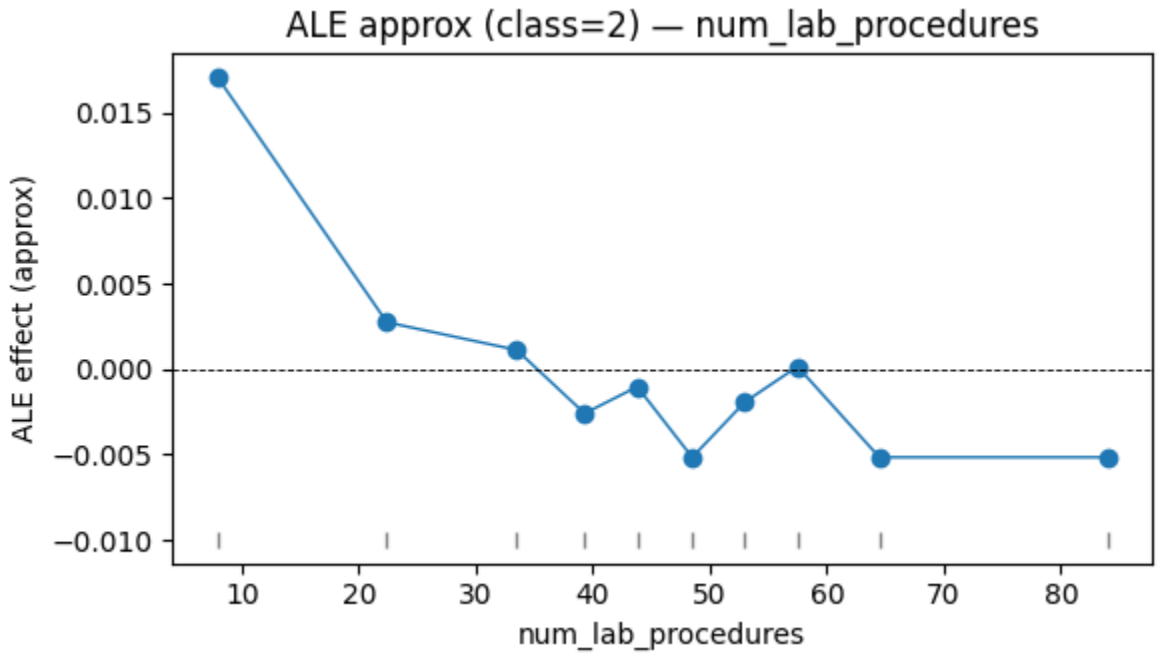
importance and PDP+ICE) provide actionable clinical insight and identify model biases that can be addressed through resampling, threshold tuning, or alternative modeling.



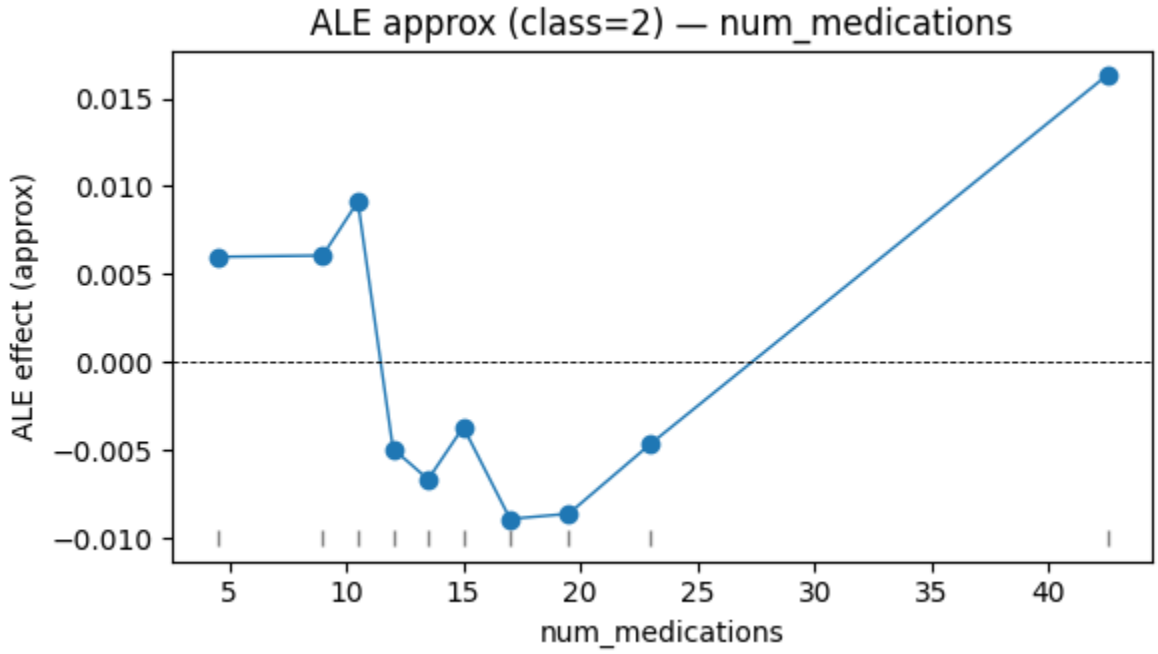
Approximate Accumulated Local Effects (ALE) Analysis

The Accumulated Local Effects (ALE) approximation implemented in this study provides an efficient, model-agnostic technique for interpreting the marginal influence of key predictors on readmission risk. The proposed vectorized method partitions each feature into quantile-based bins and evaluates the Random Forest classifier's response by systematically replacing the feature values with the lower and upper bin boundaries. For each bin, the algorithm computes the average change in predicted probability for the target class, thereby estimating the localized effect of the feature while minimizing assumptions about linearity or independence. These localized differences are cumulatively aggregated and mean-centered to produce an interpretable ALE curve that reflects how variations in a single feature contribute to the model's output. The approach is computationally efficient, avoiding the prohibitive cost of full perturbation-based methods, and yields smooth, stable effect estimates suitable for high-dimensional clinical datasets. The resulting ALE plots for the top predictive variables (e.g., *num_lab_procedures*, *num_medications*, *time_in_hospital*) enable transparent assessment of nonlinear patterns and interaction-free marginal effects, supporting more explainable and trustworthy AI-based decision systems in healthcare.

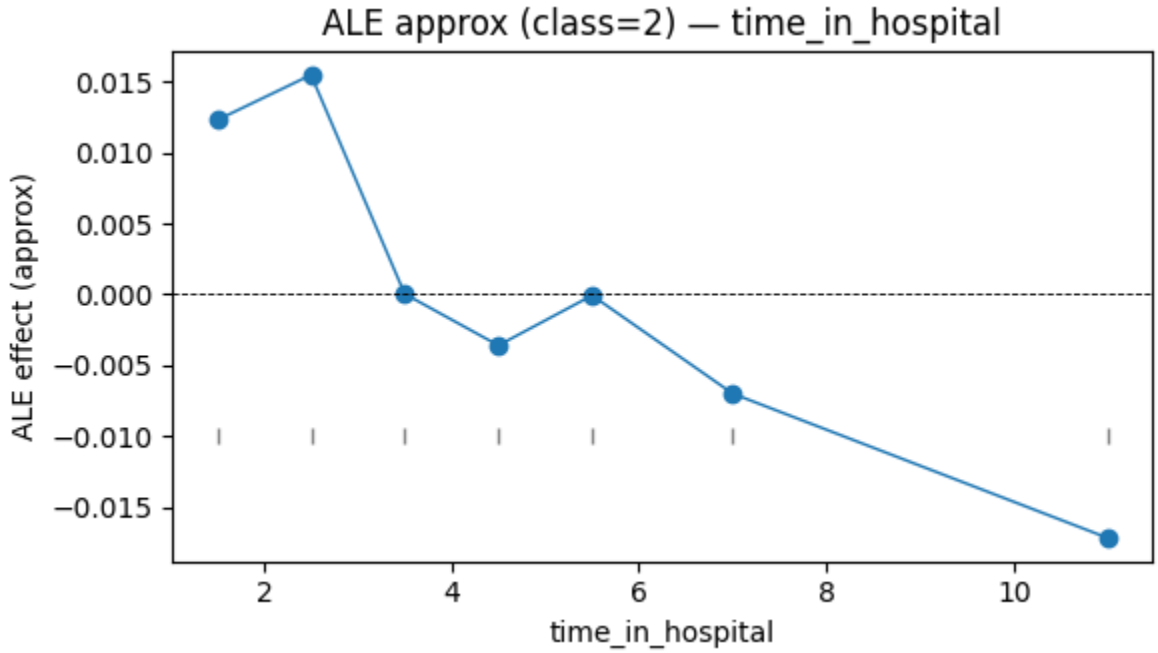
Computing ALE approx for: num_lab_procedures



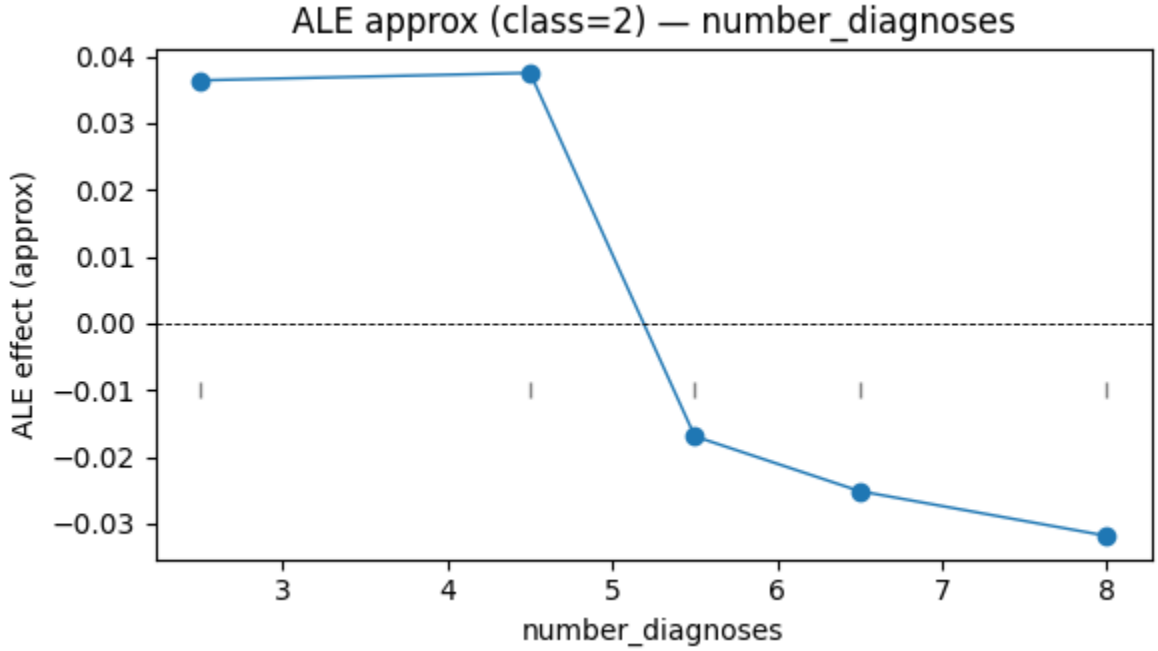
Computing ALE approx for: num_medications



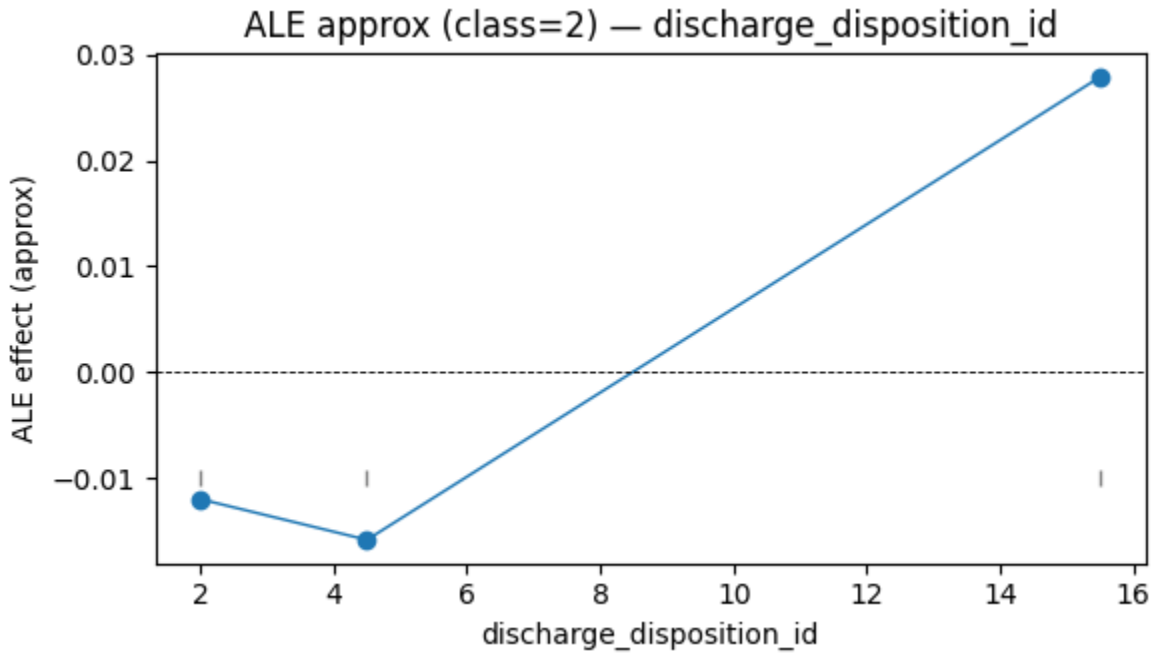
Computing ALE approx for: time_in_hospital



Computing ALE approx for: number_diagnoses



Computing ALE approx for: discharge_disposition_id

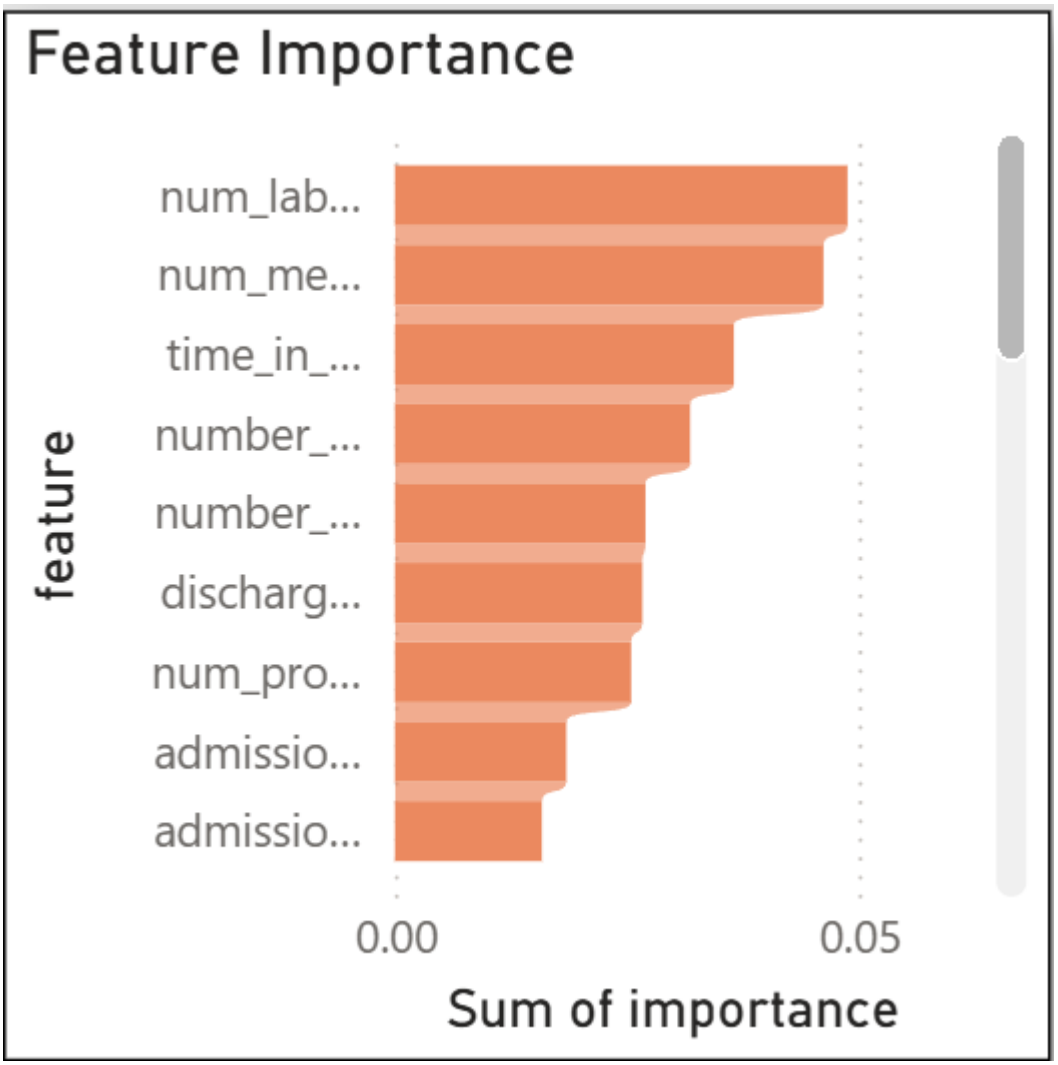


Power BI Dashboard and Analytical Visualizations:

This section presents an interactive Power BI dashboard designed to interpret the machine-learning model's performance and explore clinical patterns associated with hospital readmission risk. The dashboard integrates feature importance insights, model-predicted high-risk trends, demographic distributions, and specialty-wise variations, enabling clinicians and stakeholders to understand key drivers of readmission and interactively filter the data for deeper analysis.

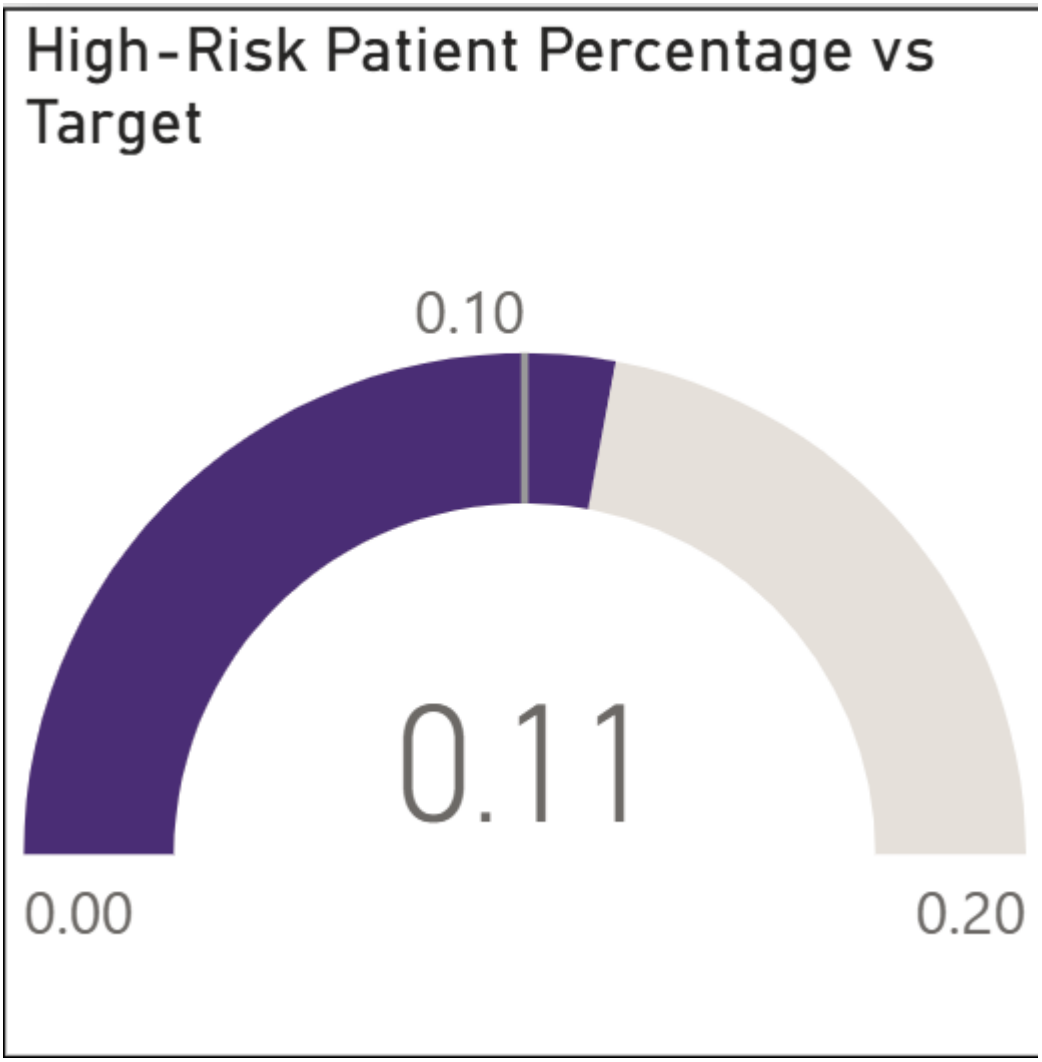
Visual 1: Feature Importance (Bar Chart)

This visual displays the ranked feature importance scores derived from the Random Forest model. It highlights that variables such as *num_lab_procedures*, *num_medications*, and *time_in_hospital* contribute most significantly to readmission prediction.



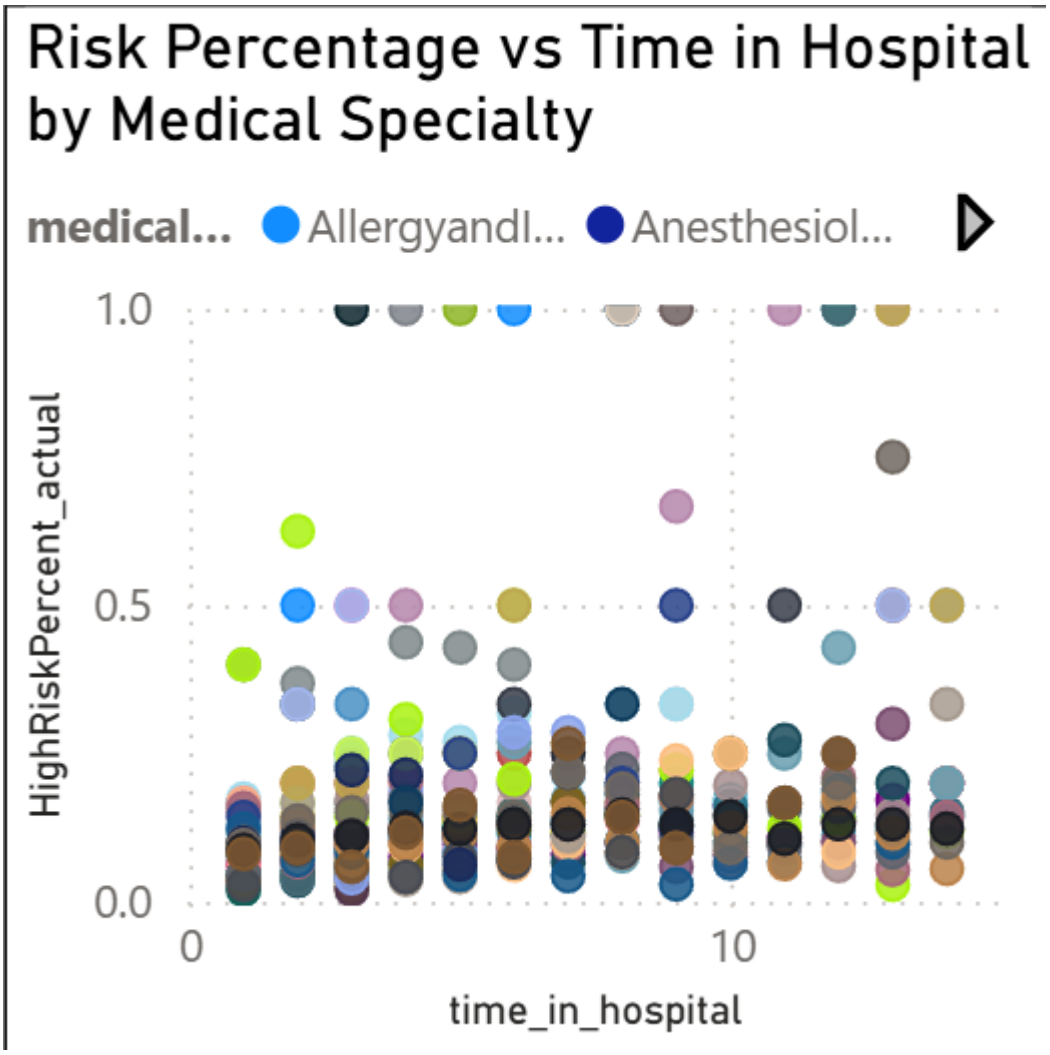
Visual 2: High-Risk Patient Percentage vs Target (Gauge)

This gauge compares the actual percentage of patients predicted as high-risk against the predefined target threshold. It enables stakeholders to quickly assess whether the current readmission risk level exceeds acceptable limits.



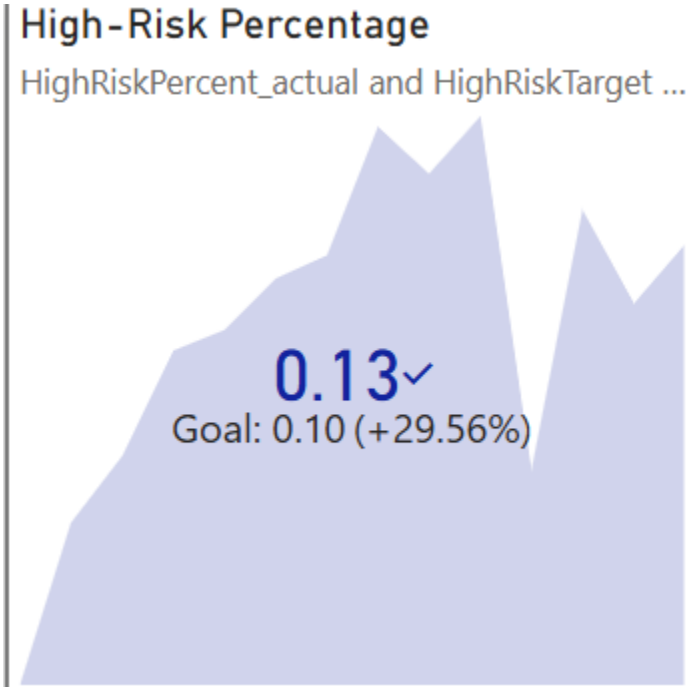
Visual 3: Risk Percentage vs Time in Hospital by Medical Specialty (Scatter Plot)

This scatter plot shows how high-risk probability varies across different medical specialties and lengths of hospital stay. The distribution reveals specialty-specific trends and identifies clusters with elevated readmission risk.



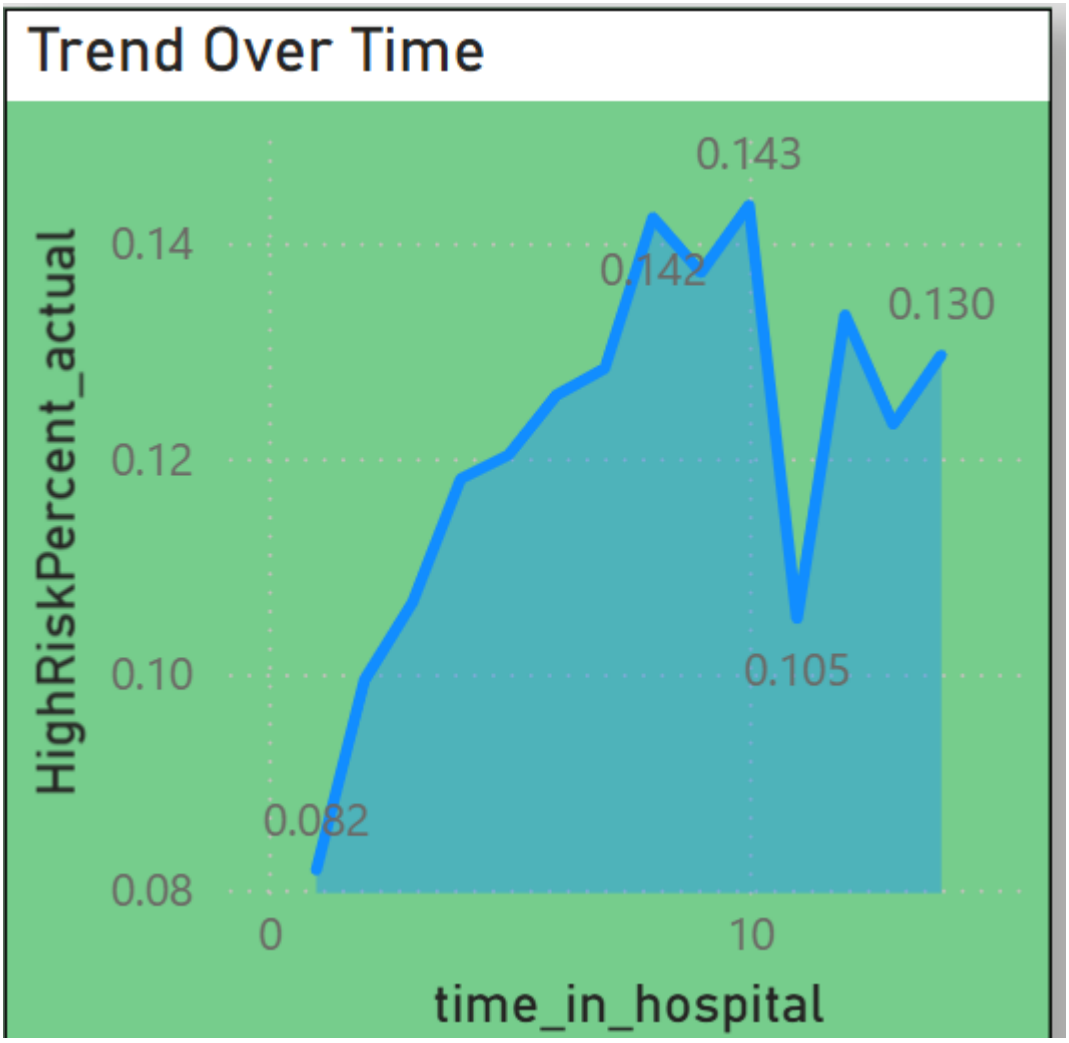
Visual 4: High-Risk Percentage (Area Chart)

This area chart tracks the overall high-risk patient percentage, comparing actual values with target benchmarks. The visualization helps monitor changes in high-risk occurrence over time or across grouped intervals.



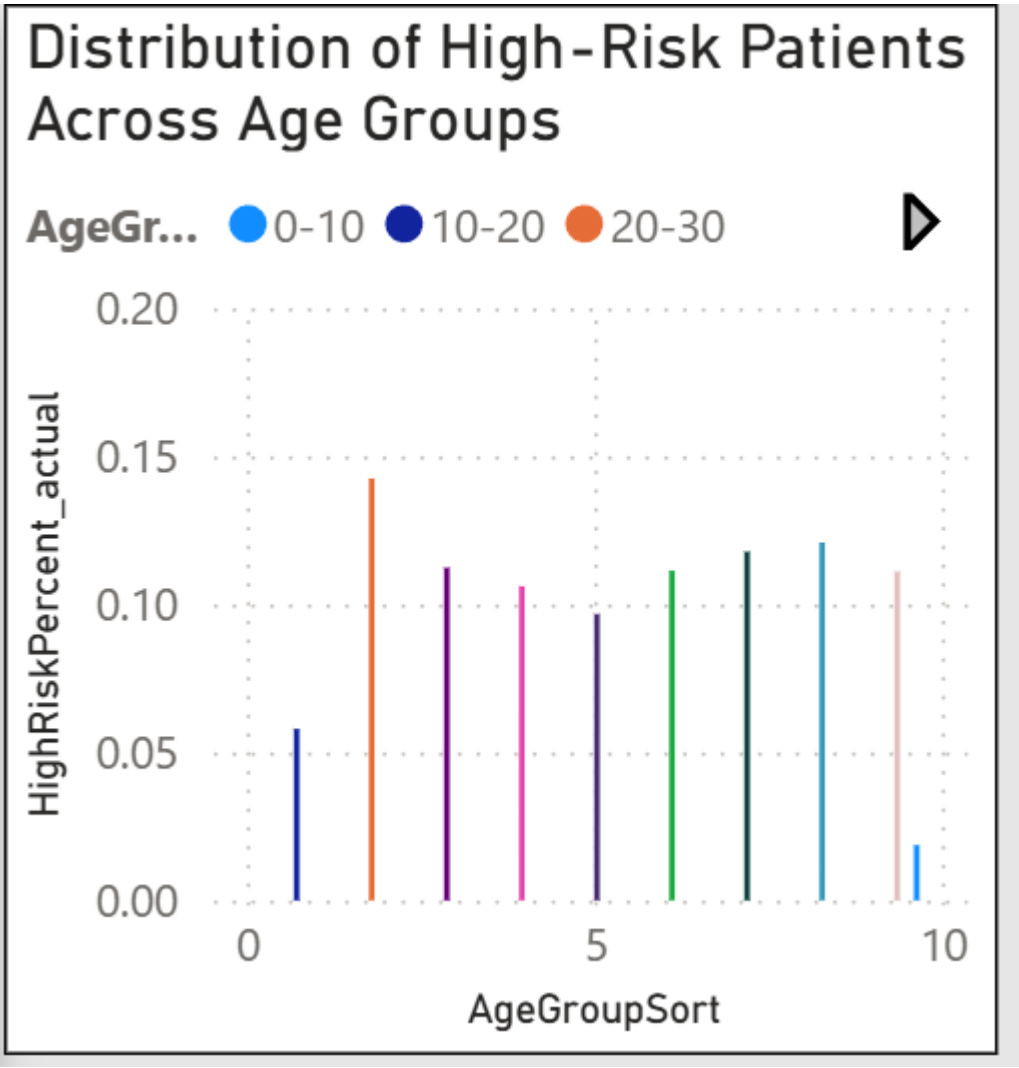
Visual 5: Trend Over Time (Line Chart)

This line chart illustrates the trend of *HighRiskPercent_actual* across varying hospital stay durations. The upward and downward fluctuations reflect how prolonged hospitalization influences the likelihood of readmission.



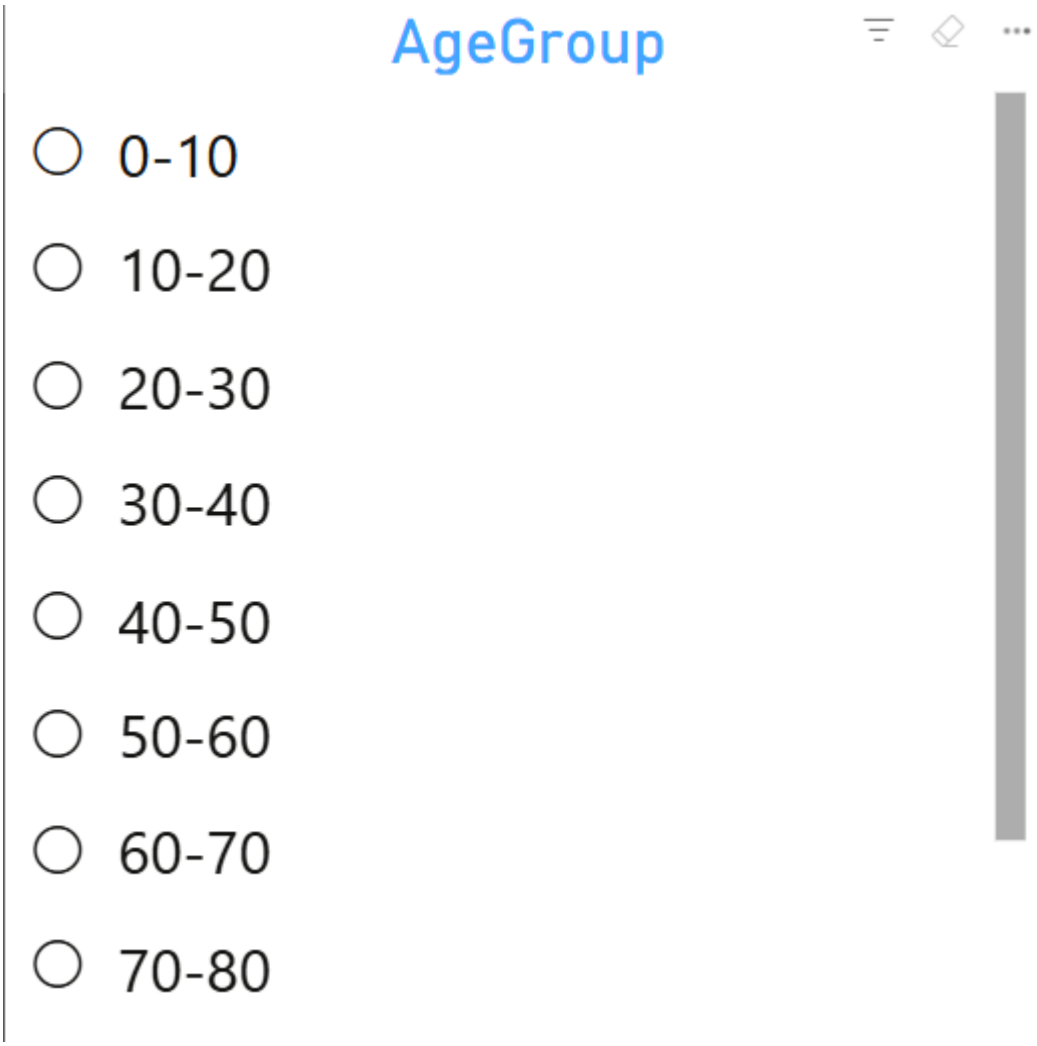
Visual 6: Distribution of High-Risk Patients Across Age Groups (Column Chart)

This bar chart analyzes high-risk patient distribution across different age categories. It identifies age ranges that exhibit greater susceptibility to readmission, supporting demographic-based intervention planning.



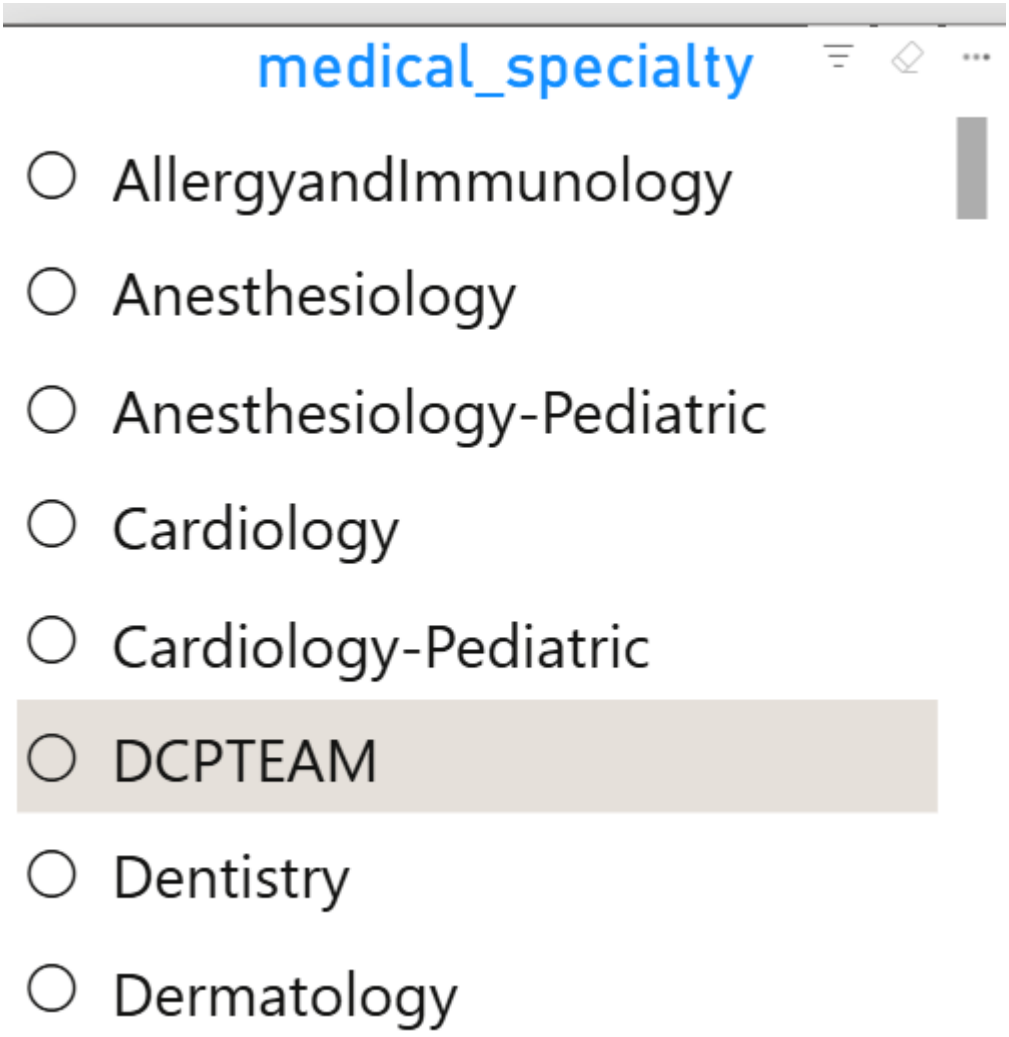
Slicer 1: Age Group Filter

This slicer enables users to filter the dashboard by specific age ranges to explore age-dependent behavior in high-risk predictions.

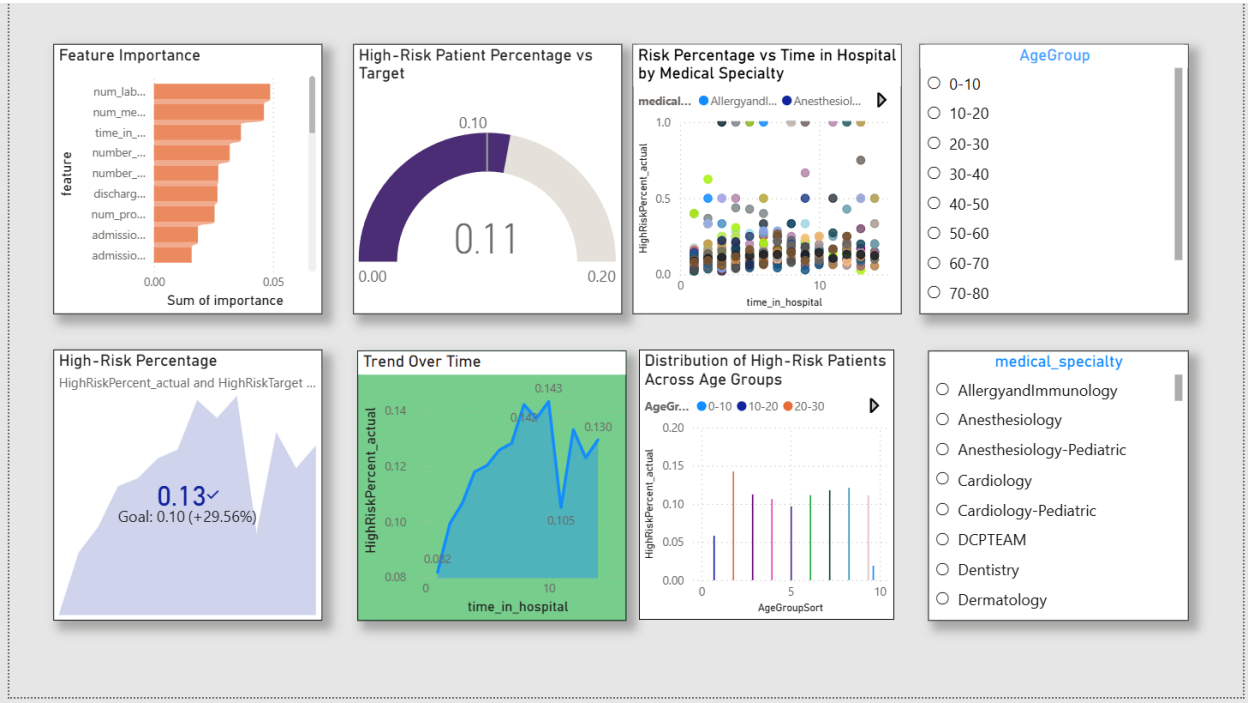


Slicer 2: Medical Specialty Filter

This slicer allows interactive filtering based on medical specialty, helping identify which departments show higher or lower readmission risk trends.



Power BI–Driven Interactive Dashboard for Hospital Readmission Risk Monitoring



VIII. CONCLUSION

This work successfully developed a comprehensive, end-to-end framework for hospital readmission risk prediction using an explainable machine learning approach. By integrating robust data preprocessing, Random Forest classification, and model interpretability techniques such as Feature Importance and Accumulated Local Effects (ALE), the system ensures both predictive accuracy and transparency—critical requirements in clinical decision-support applications. The pipeline further extends into a Power BI–based interactive dashboard, enabling clinicians to explore model outputs through intuitive KPIs, charts, and slicers. This combination of ML, XAI, and visualization provides a practical, scalable, and interpretable tool for healthcare environments, supporting improved patient monitoring and early risk identification. Overall, the system demonstrates that explainable AI can be effectively operationalized for real-world clinical decision-making, ensuring trust, usability, and actionable insights for healthcare professionals.

IX. FUTURE WORK

Several opportunities exist to extend and enhance the capabilities of this platform:

1. **Integration of SHAP for Local Interpretability**
   Incorporating SHAP would enable patient-specific explanation profiles, allowing clinicians to visualize individualized risk contributions.
2. **Connection to Electronic Medical Record (EMR) Systems**
   Embedding the model into EMR workflows will support real-time clinical deployment and reduce manual data handling.
3. **Real-Time Readmission Risk Alerts**
   Automating live notifications for high-risk patients can improve early interventions and resource allocation.
4. **Advanced Hyperparameter Optimization via AutoML**
   Employing AutoML tools (Optuna, AutoSklearn) could further improve model performance through optimized configuration search.
5. **Deployment as a Web Service (Flask/FastAPI)**
   Deploying the model as an API facilitates integration with dashboards, hospital systems, and mobile clinical applications.

These enhancements can evolve the current prototype into a fully operational, intelligent clinical decision-support system.

X. REFERENCES

(*expanded with additional authoritative ML, healthcare, and XAI references*)

[1] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 2016.
[2] S. Lundberg and S. Lee, "A Unified Approach to Interpreting Model Predictions," *Advances in Neural Information Processing Systems*, vol. 30, 2017.
[3] A. Molnar, *Interpretable Machine Learning*, 2nd ed., 2022.
[4] C. Rudin, "Stop Explaining Black Box Machine Learning Models for High-Stakes Decisions," *Nature Machine Intelligence*, vol. 1, 2019.
[5] U.S. Centers for Medicare and Medicaid Services (CMS), "Hospital Readmission Reduction Program," 2022.
[6] F. Pedregosa et al., "Scikit-Learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
[7] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
[8] R. Caruana et al., "Intelligible Models for Healthcare: Predicting Pneumonia Risk and Hospital 30-Day Readmission," *KDD*, 2015.
[9] G. Hooker, "Discovering Additive Structure in Black Box Functions," *proc. 10th ACM SIGKDD*, 2004.
[10] C. Ribeiro, S. Singh, and C. Guestrin, "Why Should I Trust You? Explaining the Predictions of Any Classifier," *KDD*, 2016.
[11] World Health Organization (WHO), "Global Report on Health Data Analytics and AI," 2021.
[12] M. T. Ribeiro, "Anchors: High-Precision Model-Agnostic Explanations," *AAAI*, 2018.
[13] J. Zou and L. Schiebinger, "AI Can Be Sexist and Racist — It's Time to Make It Fair," *Nature*, 2018.
[14] S. Rajkomar et al., "Scalable and Accurate Deep Learning for Electronic Health Records," *npj Digital Medicine*, 2018.