

Phase 6: Model Comparison - All Models

Peter Mangoro

2025-12-07

Contents

1	Introduction	1
2	Load Results	1
3	Performance Metrics Comparison	1
4	Visualization	2
4.1	Performance Metrics Comparison	2
4.2	AUC Comparison	5
4.3	ROC Curves Comparison	5
5	Interpretability Comparison	5
6	Model Selection	6
7	Key Findings	6
8	Summary	7

1 Introduction

This document presents Phase 6: Model Comparison. We compare the performance of Logistic Regression, CART, Random Forest, XGBoost, and Neural Network models, analyze their interpretability, and justify model selection.

2 Load Results

3 Performance Metrics Comparison

Table 1: Side-by-Side Performance Comparison: All Five Models

Metric	Logistic_Regression	CART	Random_Forest	XGBoost	Neural_Network
Accuracy	61.84	60.78	61.46	62.70	60.72
Precision	64.09	58.71	60.67	63.07	60.00
Recall (Sensitivity)	42.05	55.83	51.21	49.84	49.36
Specificity	79.26	65.16	70.55	74.10	70.80
F1-Score	50.78	57.23	55.54	55.68	54.16
AUC	64.81	60.50	64.82	66.25	64.16

4 Visualization

4.1 Performance Metrics Comparison

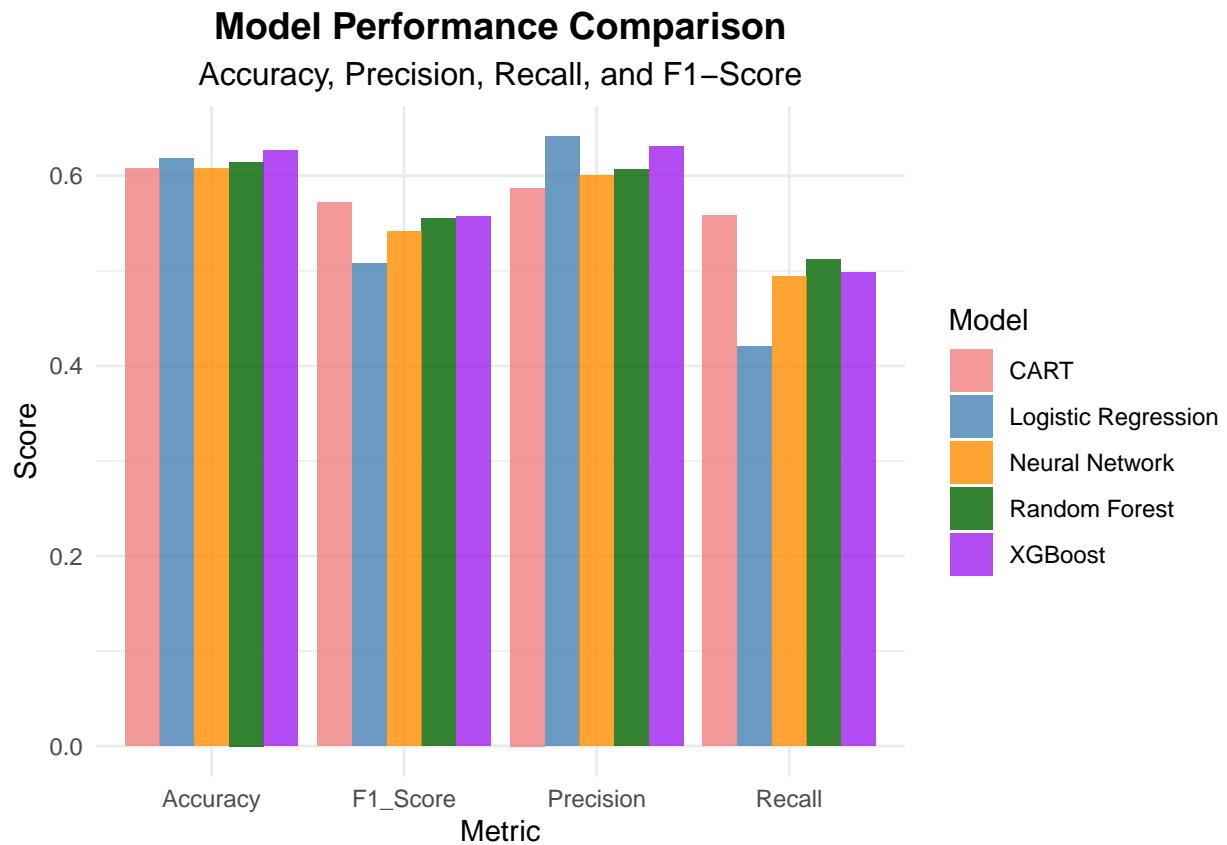


Figure 1: Model Performance Comparison

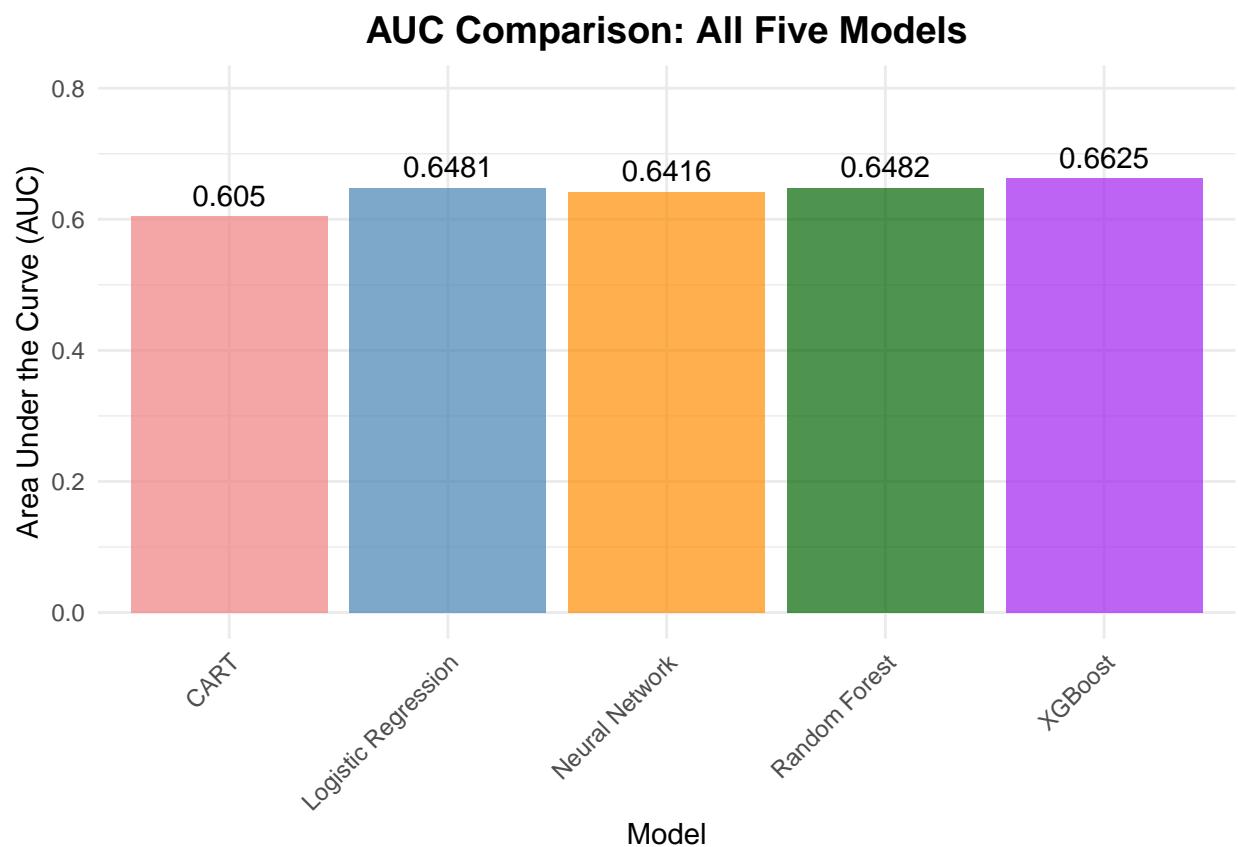


Figure 2: AUC Comparison: All Five Models

ROC Curves: All Five Models

AUC: LR=0.648, CART=0.605, RF=0.648, XGB=0.662, NN=0.642

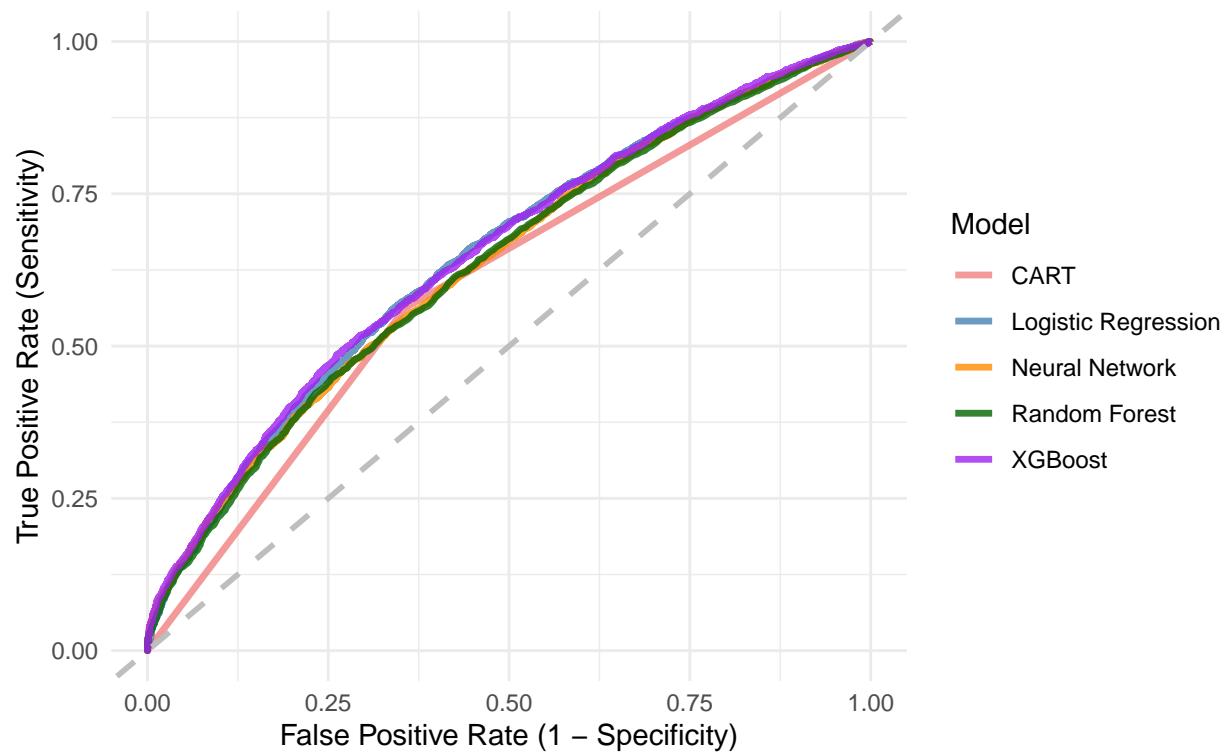


Figure 3: ROC Curves: All Five Models

4.2 AUC Comparison

4.3 ROC Curves Comparison

5 Interpretability Comparison

Logistic Regression:

- Provides coefficients and odds ratios for each variable
- Statistical significance testing (p-values)
- 29 parameters in the model
- 17 significant variables ($p < 0.05$)

CART:

- Simple decision tree with 6 variables considered
- Very interpretable: simple decision rule(s)
- Non-linear relationships captured
- Top variable: total_previous_visits (42.32% importance)

Random Forest:

- Ensemble of 500 decision trees
- Uses bootstrap sampling and feature randomization
- 17 variables considered
- Lower interpretability (ensemble effect)
- Non-linear relationships captured
- Top variable: n_lab_procedures (16.67% importance)

XGBoost:

- Gradient boosting ensemble with 200 iterations
- 17 variables considered
- Low interpretability (complex ensemble)
- Handles non-linear relationships and feature interactions
- Top variable: total_previous_visits (17.32% importance)

Neural Network:

- Feedforward network with 10 hidden units
- Very low interpretability (black box model)
- Captures complex non-linear relationships
- Requires feature scaling
- No direct variable importance (weights are not easily interpretable)

6 Model Selection

Performance Metrics Won: Logistic Regression: 2 metrics CART: 2 metrics Random Forest: 0 metrics XGBoost: 2 metrics Neural Network: 0 metrics

Recommended Model: XGBoost Reason: Highest AUC and accuracy

Table 2: Model Comparison: Detailed Justification

Criterion	Logistic_Regression	CART	Random_Forest	XGBoost	Neural_Network
Accuracy	61.84%	60.78%	61.46%	62.7%	60.72%
AUC	64.81%	60.5%	64.82%	66.25%	64.16%
Precision	64.09%	58.71%	60.67%	63.07%	60%
Recall	42.05%	55.83%	51.21%	49.84%	49.36%
F1-Score	0.5078	0.5723	0.5554	0.5568	0.5416
Interpretability	High (coefficients, odds ratios)	Very High (simple tree, easy rules)	Low (ensemble of 500 trees)	Low (gradient boosting, 200 iterations)	Very Low (black box, no direct interpretation)
Complexity	High (29 parameters)	Very Low (simple tree)	High (500 trees, complex ensemble)	High (complex ensemble, 200 trees)	Very High (10 hidden units, complex weights)
Statistical Rigor	High (p-values, hypothesis tests)	Medium (no p-values, variable importance)	Medium (variable importance, no p-values)	Low (variable importance only)	Very Low (no statistical tests, weights not interpretable)

7 Key Findings

1. Performance:

- All five models show similar performance (accuracy ~61%)
- Best AUC: XGBoost (0.662)
- Logistic Regression: 0.648, CART: 0.605, Random Forest: 0.648, XGBoost: 0.662, Neural Network: 0.642
- All models have fair to poor discrimination (AUC < 0.7)

2. Interpretability:

- CART: Simplest (6 variables, single tree)
- Logistic Regression: 29 parameters, detailed statistical insights
- Random Forest: Complex (500 trees, 17 variables)
- XGBoost: Complex (200 iterations, 17 variables)
- Neural Network: Very complex (10 hidden units, black box)
- All models identify similar key predictors

3. Key Predictors:

- Logistic Regression: n_inpatient (OR: 1.47), age groups, medical specialty
- CART: total_previous_visits (42.32% importance)
- Random Forest: n_lab_procedures (16.67% importance)
- XGBoost: total_previous_visits (17.32% importance)
- Neural Network: No direct variable importance (black box model)

8 Summary

This phase compared all five models:

- **Logistic Regression** wins 2 out of 6 performance metrics
- **CART** offers superior simplicity and interpretability
- **Random Forest** wins 0 out of 6 performance metrics
- **XGBoost** wins 2 out of 6 performance metrics
- **Neural Network** wins 0 out of 6 performance metrics
- **Recommended:** XGBoost (Highest AUC and accuracy)
- All tree-based models identify **total_previous_visits** as a key predictor
- Ensemble methods (RF, XGBoost) show similar performance
- Neural Network provides alternative approach but with lowest interpretability