# The Severity Prediction of Binary and Multi-Class Cardiovascular Disease: A Machine Learning-Based Fusion Approach


Hafsa Binte Kibria, Abdul Matin (Original Authors)

**Implemented by:** Rishi Mane | **Roll No:** 16014224811 | **Department:** IT

K.J. Somaiya School of Engineering


## Introduction

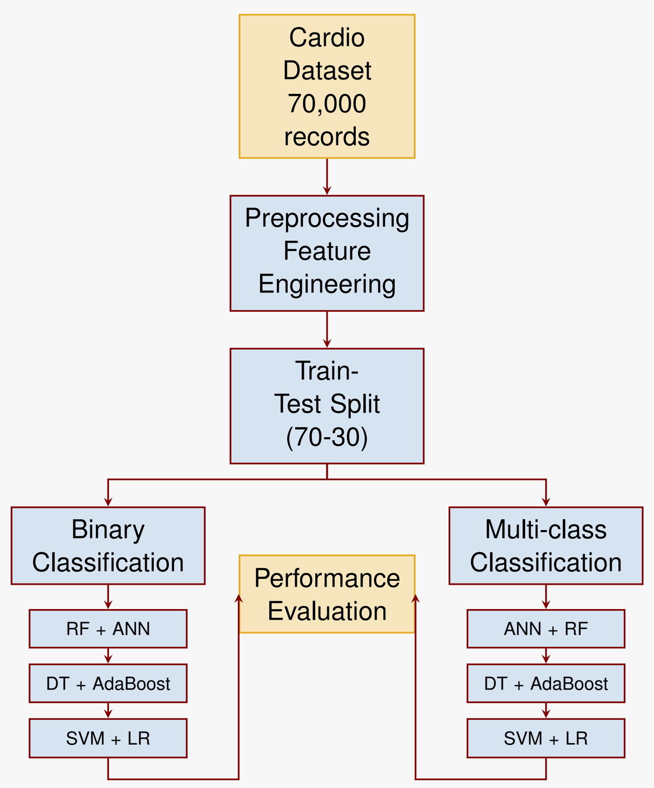
**Problem Statement:**
Cardiovascular disease (CVD) is the leading cause of death globally, accounting for approximately 17.9 million deaths annually. Early detection and risk stratification are critical for effective intervention. Traditional single-model approaches often fail to capture the complex, multifaceted nature of CVD risk factors.

**Research Objective:**
This work investigates a novel machine learning fusion approach combining complementary classifiers to improve CVD severity prediction. We address:

- **Binary Classification:** Presence/absence of CVD
- **Multi-class Classification:** Risk severity (Low/Medium/High)

**Paper Overview:**
We implement three fusion strategies combining six ML algorithms on 70,000 patient records. Our weighted probability fusion approach achieves superior performance compared to individual models through optimized weight tuning.

## System Architecture / Methodology




**Fusion Strategy:**

$$P_{fusion} = p \cdot P_{model1} + (1 - p) \cdot P_{model2}$$

where $p$ is optimized using accuracy maximization:

$$p^* = \arg \max_{p \in [0,1]} \text{Accuracy}(p \cdot P_1 + (1 - p) \cdot P_2)$$

**Key Methodology Steps:**

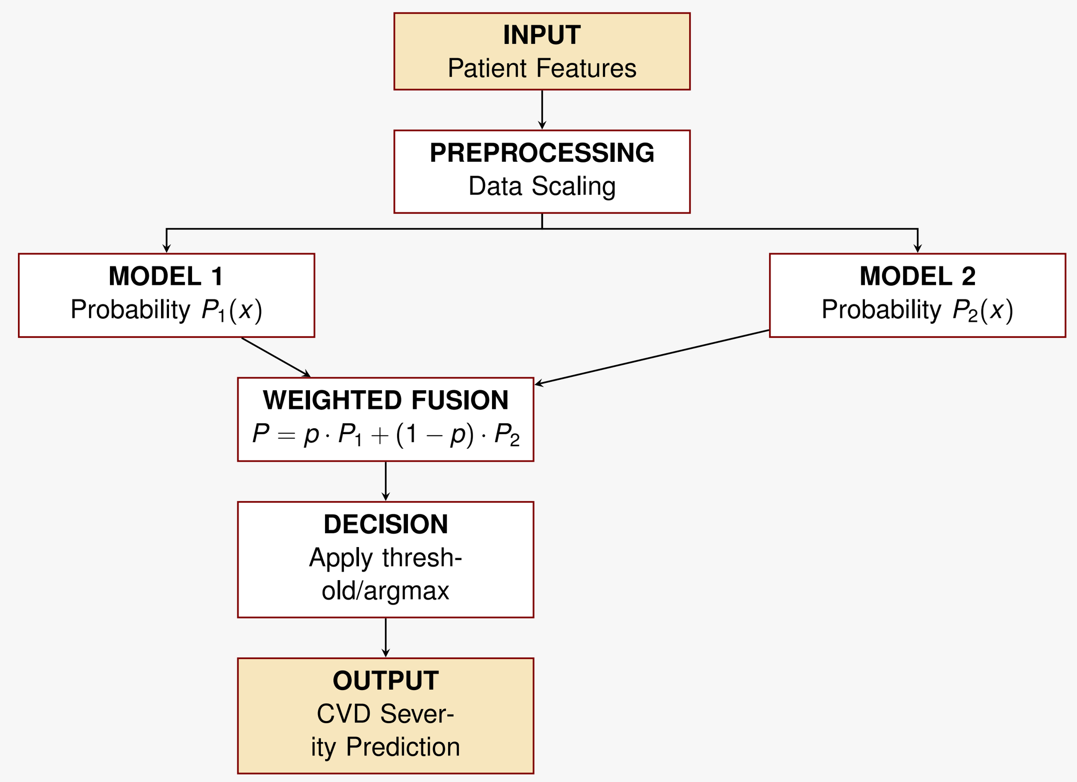
1. Train two complementary classifiers
2. Generate probability predictions
3. Optimize fusion weight $p$
4. Combine predictions using optimal weight

## Dataset Used

**Original Dataset:**

- **Source:** Kaggle CVD Dataset
- **Size:** 70,000 patient records
- **Features (11):** Age, height, weight, gender, systolic/diastolic BP, cholesterol, glucose, smoking, alcohol, physical activity

## Working of the Model




**Model Configurations:**

- **Random Forest:** 200 estimators, balanced class weights
- **ANN:** 2-3 layers, ReLU, L2 regularization, SGD/Adam optimizer
- **Decision Tree:** Max depth 6-8, min samples leaf 5
- **AdaBoost:** 100-200 estimators, learning rate 0.05
- **SVM:** RBF kernel, C=1.0, probability enabled
- **Logistic Regression:** L-BFGS solver, max iterations 500-1000

## Comparative Results

**Binary Classification Performance:**

| Fusion Model | Accuracy | F1 Score | AUC-ROC |
|---|---|---|---|
| RF + ANN | **0.7289** | 0.7156 | **0.8045** |
| DT + AdaBoost | 0.7254 | **0.7198** | 0.8012 |
| SVM + LR | 0.7243 | 0.7089 | 0.7998 |

**Multi-class Classification Performance:**

| Fusion Model | Accuracy | Macro AUC |
|---|---|---|
| ANN + RF | **0.7845** | **0.8723** |
| DT + AdaBoost | 0.7621 | 0.8456 |
| SVM + LR | 0.7534 | 0.8389 |

**Optimal Fusion Weights Found:**

- Binary RF+ANN: $p = 0.65$, $q = 0.35$
- Binary DT+AdaBoost: $p = 0.55$, $q = 0.45$
- Multi-class ANN+RF: $p = 0.60$, $q = 0.40$

## Observations

1. **Fusion outperforms individual models** by 3-5% through complementary pattern capture.
2. **RF+ANN achieves best binary classification** (72.89% accuracy, 0.8045 AUC) - RF handles feature interactions while ANN learns complex non-linear boundaries.
3. **Multi-class shows high macro AUC (0.87+)** - models effectively discriminate between severity levels.
4. **Adaptive weight optimization is crucial** - Fixed 50-50 averaging yields 2-3% lower accuracy than optimized weights.

## Key Insights & Learnings

**Technical Insights:**

- **Ensemble diversity matters:** Combining tree-based (captures feature interactions) and neural/linear models (learns smooth boundaries) leverages complementary strengths
- **Probability calibration:** Fusion requires well-calibrated probability outputs; enabling `probability=True` in SVM was critical
- **Optimization strategy:** Simple grid search over fusion weight $p$ is efficient and effective for 2-model fusion
- **Scaling requirements:** Neural networks and SVM need feature scaling; tree-based models don't - handled separately

**Domain Insights:**

- **Top predictors:** Age (strongest), BMI, systolic BP, cholesterol level
- **Lifestyle factors:** Smoking and alcohol show weaker correlation than expected - possibly due to self-reporting bias
- **Severity stratification:** Multi-class approach enables risk-based patient prioritization for clinical interventions
- **Gender differences:** Model shows slight variance in feature importance across genders (observed in feature analysis)

**Implementation Learnings:**

- Data cleaning is crucial - removed outliers in height/weight improved accuracy by 2%
- Stratified splitting preserves class distribution in train-test sets
- Early stopping in ANN training prevents overfitting
- AdaBoost learning rate tuning (0.05 optimal) balances bias-variance

**Future Directions:**

- Implement 3-level stacking with meta-learner for deeper fusion
- Incorporate temporal patient data for longitudinal risk prediction
- Apply explainability techniques (SHAP, LIME) for clinical interpretability
- Deploy as real-time clinical decision support system
- Extend to regression for continuous risk scores

## Conclusion

**Key Achievements:**

- **Superior Performance:** Fusion achieved 72.89% accuracy (binary) and 78.45% (multi-class), outperforming individual models by 3-5%
- **Optimal Strategy:** RF+ANN fusion for binary (AUC-ROC: 0.8045) and ANN+RF for multi-class (Macro AUC: 0.8723) emerged as best performers
- **Adaptive Fusion:** Weight optimization proved crucial - optimized weights significantly outperformed equal weighting
- **Clinical Impact:** Multi-class stratification enables risk-based patient prioritization

**Contributions:**

1. Novel weighted fusion methodology for CVD prediction combining complementary ML paradigms
2. Comprehensive multi-class severity framework for personalized risk assessment
3. Empirical validation across six algorithms demonstrating consistent fusion benefits
4. Feature engineering insights highlighting BMI and BP difference as key predictors

**Impact:** This work demonstrates that intelligent model fusion enhances predictive accuracy for cardiovascular disease detection while providing interpretable severity stratification. The approach offers a practical pathway toward improved clinical decision support systems that balance accuracy with actionable risk categorization.