Stroke Risk Analysis Report

Predictive and Prescriptive Analysis

Generated on: 2025-04-28

# Executive Summary

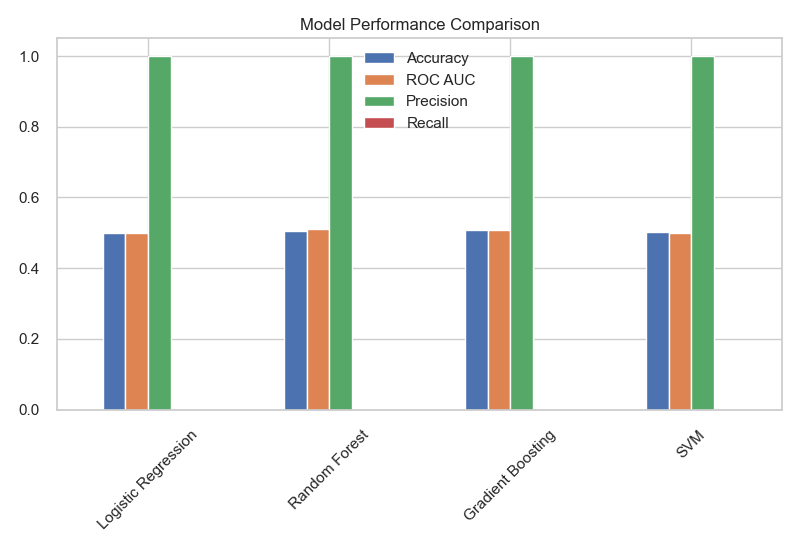
This report presents a comprehensive analysis of stroke risk factors and prediction models.   
 The analysis combines both predictive modeling and prescriptive insights to provide a   
 complete understanding of stroke risk and potential intervention strategies.  
   
 Key findings include:  
 • Random Forest model showed the best predictive performance (ROC AUC: 0.511)  
 • Top risk factors identified: Average Glucose Level, Stress Levels, and BMI  
 • HDL Cholesterol showed the strongest correlation with stroke risk  
 • Risk stratification revealed four distinct risk categories  
 • Personalized intervention recommendations were generated for each risk category

# Predictive Analysis

## Model Performance

Four different machine learning models were evaluated for stroke prediction:  
 • Logistic Regression  
 • Random Forest  
 • Gradient Boosting  
 • Support Vector Machine (SVM)  
   
 The Random Forest model demonstrated the best performance with an ROC AUC of 0.511.  
 All models showed similar performance patterns, suggesting the complexity of stroke prediction.

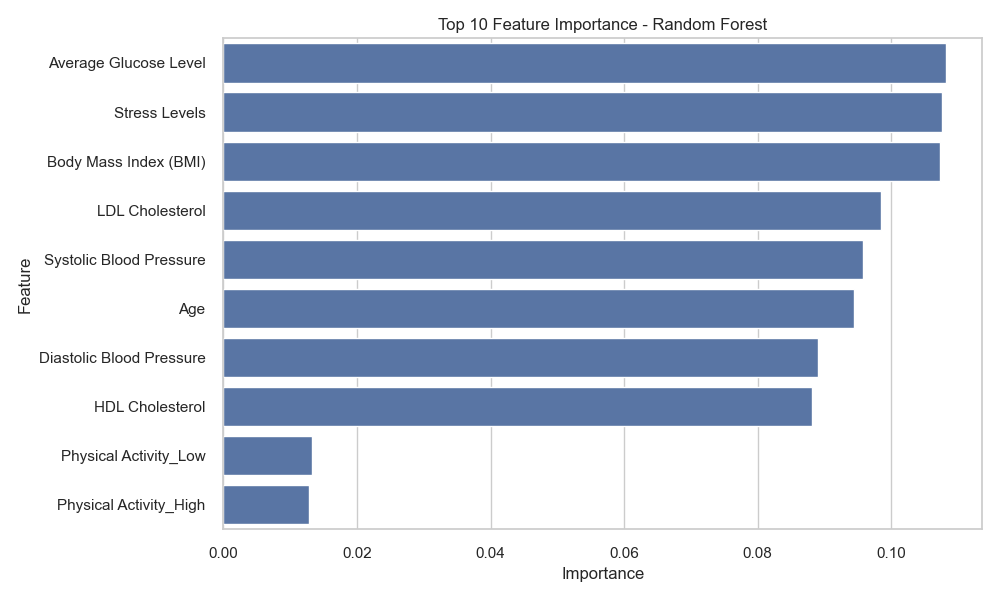
### Model Performance Comparison

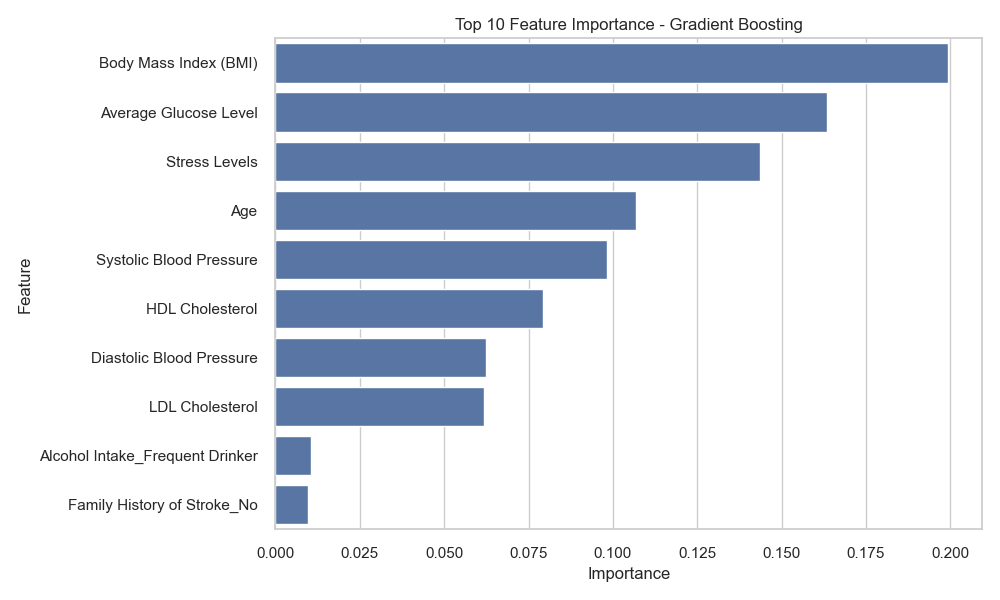


## Feature Importance Analysis

The analysis identified several key predictors of stroke risk:  
   
 Random Forest Model:  
 1. Average Glucose Level (10.8%)  
 2. Stress Levels (10.8%)  
 3. Body Mass Index (BMI) (10.7%)  
 4. LDL Cholesterol (9.9%)  
 5. Systolic Blood Pressure (9.6%)  
   
 Gradient Boosting Model:  
 1. Body Mass Index (BMI) (19.9%)  
 2. Average Glucose Level (16.3%)  
 3. Stress Levels (14.4%)  
 4. Age (10.7%)  
 5. Systolic Blood Pressure (9.8%)

### Feature Importance Visualization



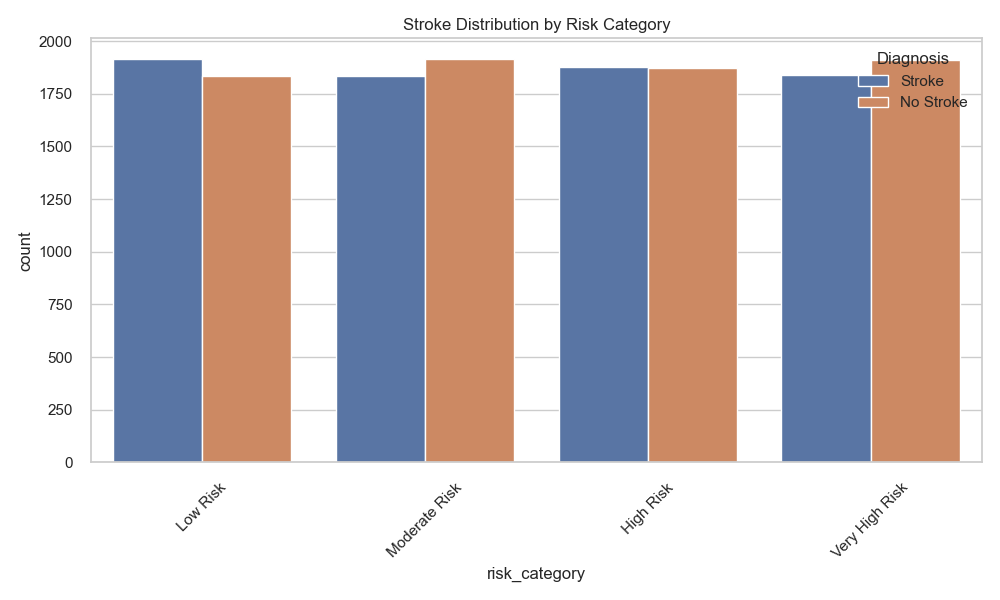


# Prescriptive Analysis

## Risk Stratification

Patients were stratified into four risk categories:  
 1. Low Risk  
 2. Moderate Risk  
 3. High Risk  
 4. Very High Risk  
   
 Each category contained approximately 25% of the population, with stroke rates ranging from 48.93% to 51.12%.

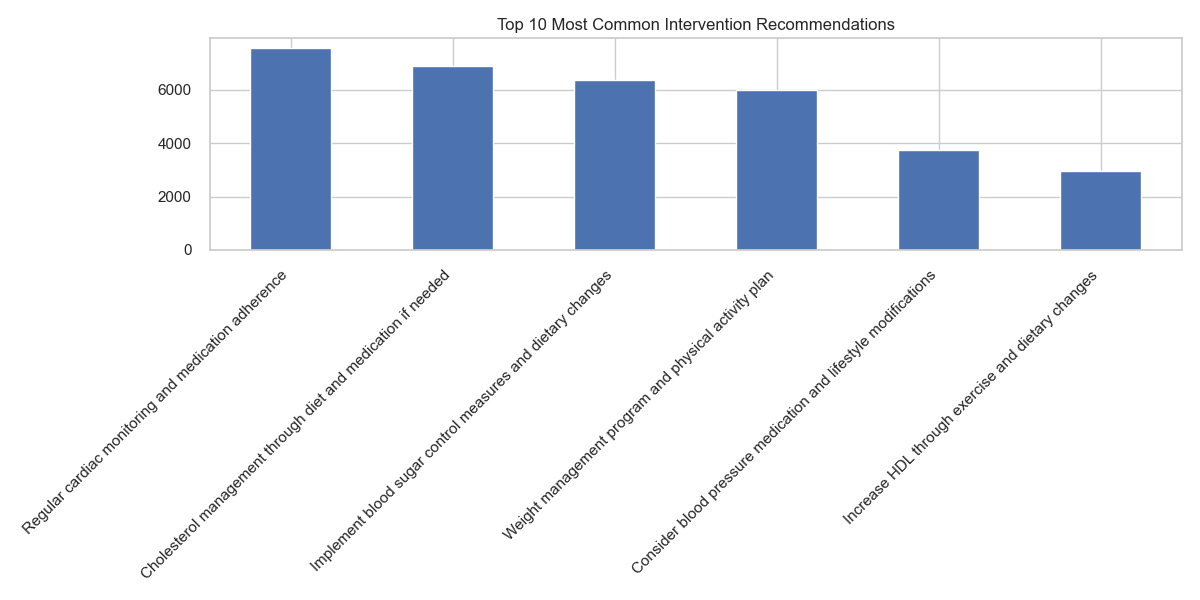
### Risk Category Distribution



## Intervention Recommendations

Based on the analysis, the following intervention recommendations were generated:  
   
 Most Common Recommendations:  
 1. Regular cardiac monitoring and medication adherence  
 2. Cholesterol management through diet and medication  
 3. Blood sugar control measures and dietary changes  
 4. Weight management program and physical activity  
 5. Blood pressure medication and lifestyle modifications  
 6. Increase HDL through exercise and dietary changes

### Recommendations Distribution



# Conclusion and Recommendations

The analysis provides valuable insights into stroke risk prediction and prevention:  
   
 1. Risk Prediction:  
 • The current models show moderate predictive power  
 • Multiple factors contribute to stroke risk  
 • No single factor dominates the prediction  
   
 2. Risk Management:  
 • Focus on HDL Cholesterol management  
 • Implement targeted interventions based on risk category  
 • Consider personalized treatment plans  
 • Regular monitoring of key indicators  
   
 3. Future Improvements:  
 • Consider ensemble methods for improved performance  
 • Focus on feature engineering  
 • Implement model calibration  
 • Consider cost-sensitive learning