

# High-School Student Performance Prediction

## 1. Overview of Machine Learning Models

a neural network and a support vector machine (SVM). These models are aimed at predicting high school students' performance in mathematics based on a publicly available dataset.

### 1.1 Neural Network Design

- Structure:
  - Input Layer: Number of neurons equal to the number of features.
  - Hidden Layers: Three layers with 128, 64, and 32 neurons, respectively.
  - Activation Function: ReLU (Rectified Linear Unit) for all layers.
  - Dropout: 30% dropout applied to the first two hidden layers.
  - Output Layer: One neuron for regression output.
  - Optimizer: Adam.
  - Loss Function: Mean Squared Error (MSE).
  - Evaluation Metric: Mean Absolute Error (MAE).
  - Initial Weights: Initialized using the Xavier initialization

### 1.2 Support Vector Machine (SVM) Design

- Kernel: Linear kernel.
- Gamma: Scale (default setting).
- Degree: 1.
- Noise Reduction Parameter: The regularization parameter (C) is used to balance the trade-off between achieving a low error on the training data and minimizing model complexity to reduce noise. A smaller C values prioritize simplicity (higher noise tolerance), while larger values focus on accuracy (less noise tolerance).
- Evaluation Metric: Mean Absolute Error (MAE).

## 2. Experimental Design and Results Analysis

### 2.1 Cross-Validation Approach

- Method: 10-Fold Cross-Validation.
- Epochs: 5 epochs for the neural network model.
- Error Metric: Proportion of misclassifications was calculated after each epoch.

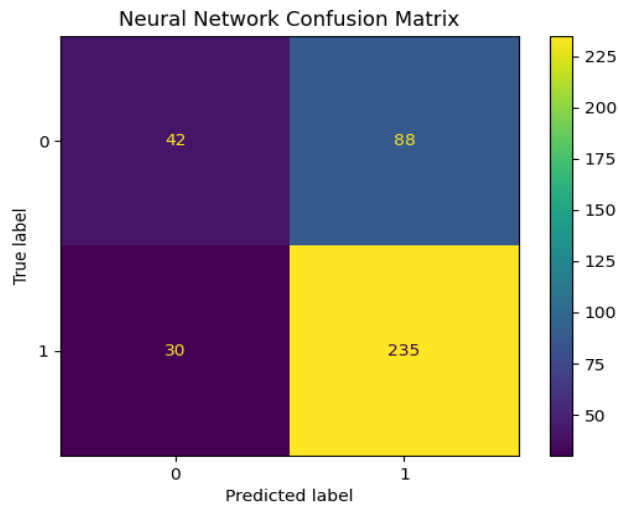
## 2.2 Performance Metrics

### Neural Network:

- Average Test Error: 29.86%

- Confusion Matrix:

```
[[ 42 88]  
 [ 30 235]]
```

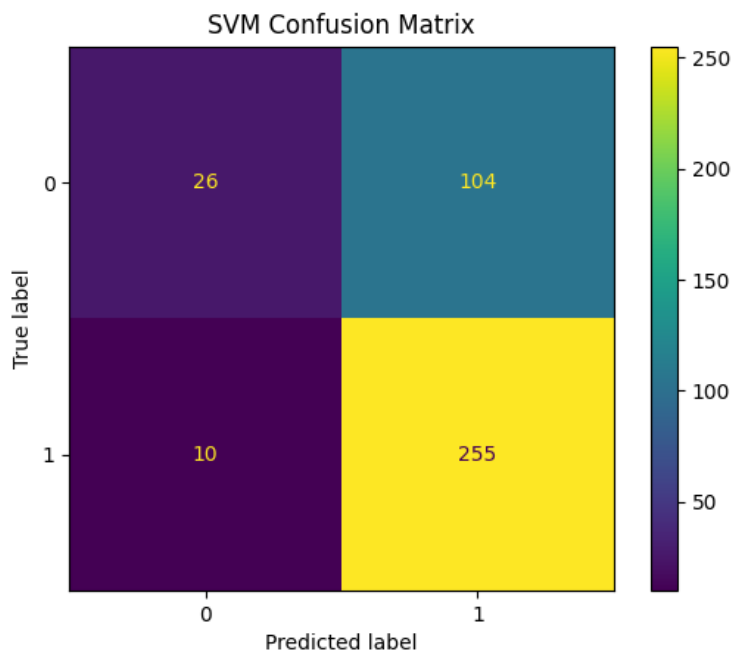


### SVM:

- Average Test Error: 28.83%

- Confusion Matrix:

```
[[ 26 104]  
 [ 10 255]]
```



2.3 Comparative Metrics Across Folds

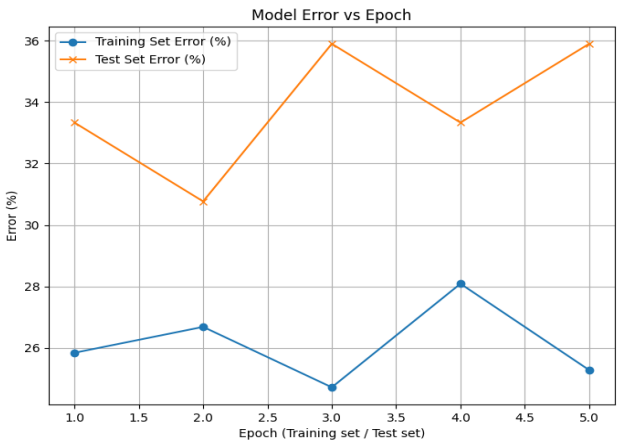
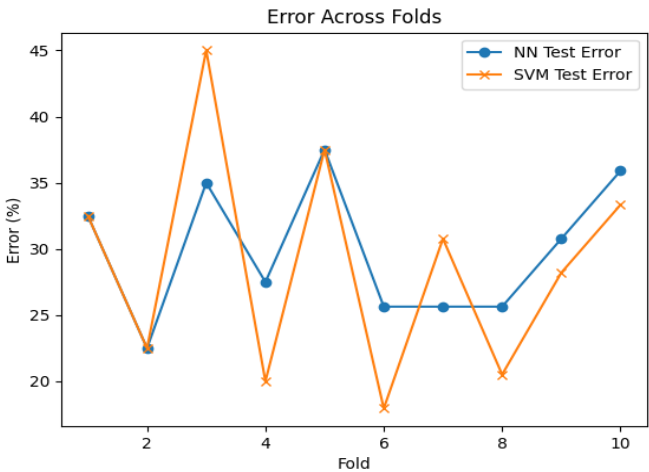
Neural Network:

Fold	Accuracy	Precision	Recall	F1-Score	
-----	-----	-----	-----	-----	
1	68%	66%	100%	79%	
2	78%	83%	86%	84%	
3	65%	61%	90%	73%	
4	72%	81%	83%	82%	
5	62%	62%	91%	74%	
6	74%	84%	84%	84%	
7	74%	74%	96%	84%	
8	74%	81%	86%	83%	
9	69%	74%	85%	79%	
10	64%	66%	88%	75%	

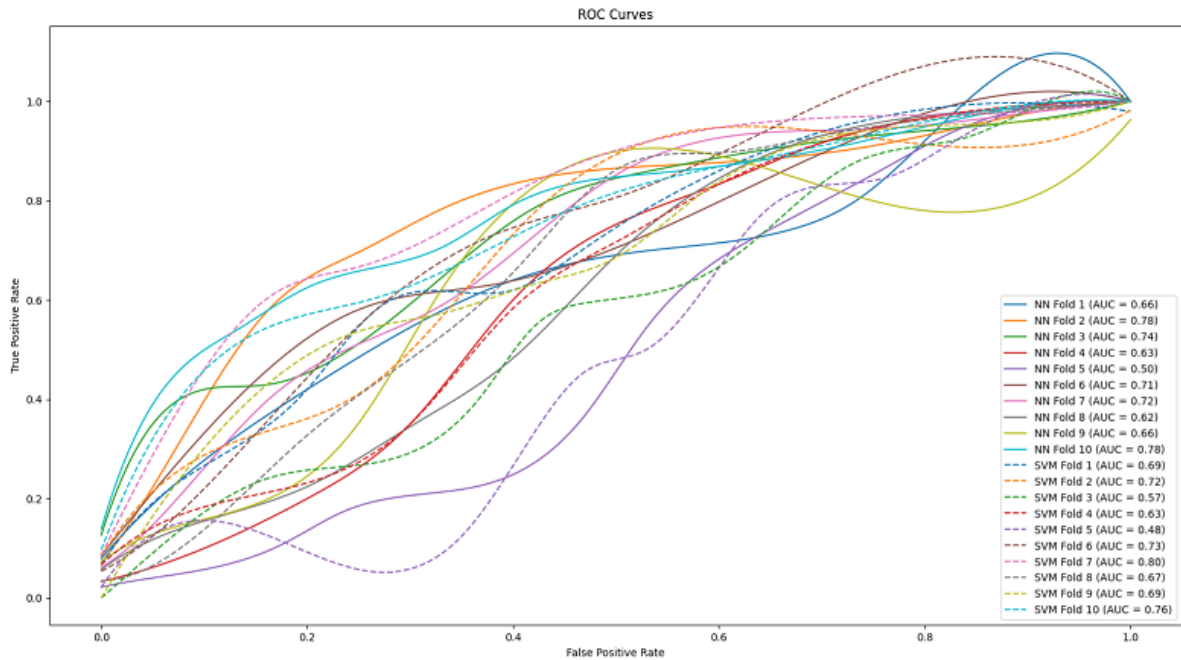
SVM:

Fold	Accuracy	Precision	Recall	F1-Score	
-----	-----	-----	-----	-----	
1	68%	67%	96%	79%	
2	78%	77%	96%	86%	
3	55%	54%	95%	69%	
4	80%	79%	100%	88%	
5	62%	61%	100%	75%	
6	82%	85%	94%	89%	
7	69%	70%	96%	81%	
8	79%	80%	97%	88%	
9	72%	71%	100%	83%	
10	67%	68%	88%	76%	

Graphs:



## Roc:



## 3. Performance Comparison and Recommendations

### 3.1 Comparison of Average Metrics

Metric	Neural Network	SVM
Accuracy (%)	70.00	71.20
Precision (%)	73.20	71.70
Recall (%)	90.90	95.80
F1-Score (%)	79.80	81.40

### 3.2 Error Rates:

- Neural Network: Average test error (29.86%).
- SVM: Lower average test error (28.83%).

### 3.3 ROC Analysis:

Both models demonstrated acceptable ROC characteristics, with SVM slightly outperforming the neural network.

#### 3.4 Recommendation:

Based on the performance metrics and lower error rate, SVM is the recommended model for predicting high school student performance in mathematics.

#### 4. Summary and Conclusion

The study implemented and compared two machine learning models for predicting student performance. While both models displayed satisfactory results, SVM consistently outperformed the neural network in terms of accuracy, recall, and F1-score. The recommendation is to use the SVM model for future predictions.