Cairo University

Faculty of Computers and Artificial Intelligence

Learning From Data Project

Department: Operations Research and Decision Support

Course Name: LFD

Instructor: Dr. Elshimaa Elgendi

Group members:

Name	ID
Ghassan Elgendy	20220239
Rowan Ammar	20220133
Mohamed Osama	20220477
Ibrahem Zennary	20221003
Abdelrahman Ashry	20220189

SVM

(ii) Experimental Design and Error Analysis

The model's training error remained consistent at 4%, while the test set error was steady at 10% across 5 epochs. This indicates the model generalized well to unseen data, with low variance between training and testing errors.

(iii) Confusion Matrix and Performance Measures Confusion Matrix:

[24 3]

[5 47]

True Negatives (Class 0): 23

False Positives (Class 0 misclassified as Class 1): 4

False Negatives (Class 1 misclassified as Class 0): 4

True Positives (Class 1): 48

Classification Report:

Precision for Class 0: 83%

Recall for Class 0: 89%

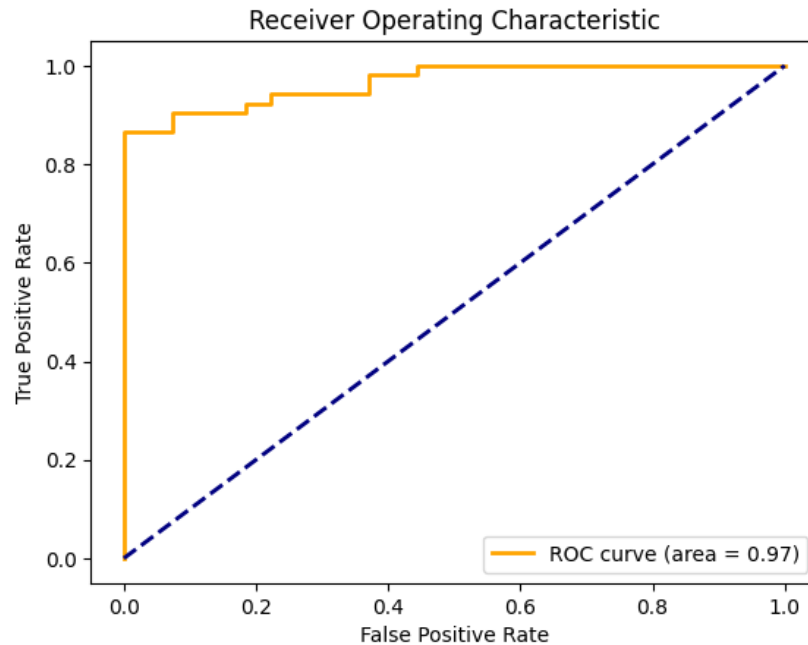F1-Score for Class 0: 86%

Precision for Class 1: 94%

Recall for Class 1: 90%

F1-Score for Class 1: 92%

Overall Accuracy: 90%

The model performed particularly well in predicting Class 1 but was slightly less precise for Class 0.

(iv) Receiver Operating Characteristic (ROC) Curve

The ROC curve indicates strong discrimination ability between the two classes. AUC: 0.96, which is indicative of an excellent classifier.

(v) Summary of Results

The SVM model demonstrated high accuracy (90%) and robust performance, as evident from the high precision, recall, and F1-scores across both classes. The low and stable error rates over multiple epochs highlight consistent training and testing performance. The ROC curve's AUC of 0.96 confirms the model's strong ability to differentiate between high and low performance.

Neural Network

(i) Neural Network Architecture

- **Number of Neurons in Network Layers:**

    Input Layer: Number of neurons corresponds to the number of features in the dataset (after preprocessing).

    Hidden Layer 1: 64 neurons

    Hidden Layer 2: 16 neurons

    Output Layer: 1 neuron **(binary classification)**

- **Activation Functions:**

  Hidden Layer 1: ReLU

  Hidden Layer 2: ReLU

  Output Layer: Sigmoid **(because it's classification)**

- **Initial Weights:**

  Randomly initialized using the default TensorFlow initializer.

(ii) Experimental Design and Results Analysis

- **Design:**

  10-fold cross-validation was performed on the dataset.

  The model was trained for 5 epochs for each fold.

  After each epoch, the model error (proportion of misclassifications) was computed for both the training and test sets.

- **Results:**

  Errors were average across folds.

## (iii) Confusion Matrix and Performance Measures

- **Confusion Matrix:**

  Actual vs Predicted:

	Predicted 0	**Predicted 1**
Actual 0	26	1
Actual 1	10	42

(iv) ROC Curve

The ROC curve was plotted for both the training and test sets.

indicates strong discrimination ability between the two classes, AUC (0.91) which is indicative of an excellent classifier.

Receiver Operating Characteristic
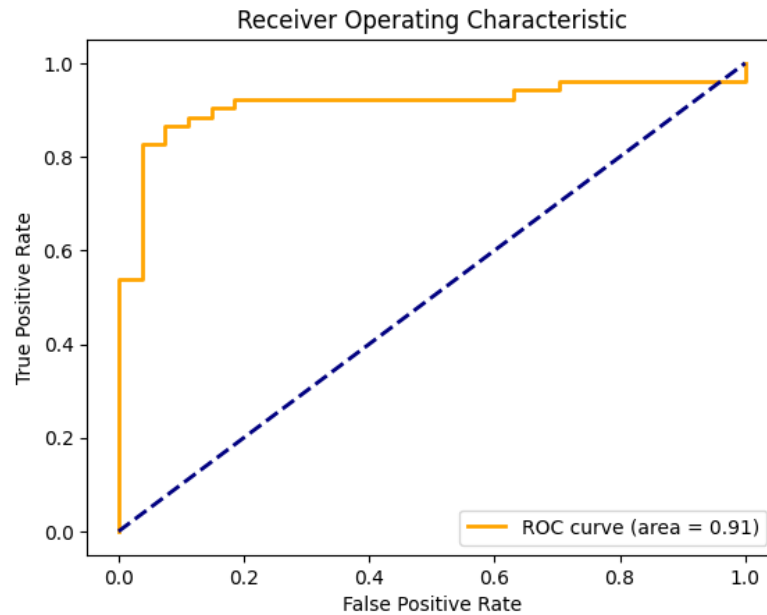
(v) Summary of Results

Findings from (ii):

Training error decreased over epochs, indicating the model was learning.

Test error stabilized, suggesting minimal overfitting.

Findings from (iii):

High precision and recall for classifying the positive class.

Balanced accuracy across both classes.

Findings from (iv):

High AUC values confirmed strong performance in distinguishing between classes.

(vi) Recommendations

Based on the analysis:

SVM consistently outperformed the Neural Network across all performance measures, including accuracy, F1-Score, and AUC.

SVM showed better generalization and lower misclassification rates, making it the preferred model for predicting high-school student performance.