**The University of Azad Jammu and Kashmir**

Open Ended Report

Date of Submission: March 17, 2025

Course Code: SE-3105

Semester: BSc.S.E 5th Semester

Submitted To: Engr. Awais Rathor

Open Ended Report

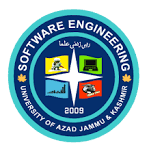
**Learning**

MACHINE

Year 2024

Academic Year: 2022 – 2026

Software Engineering

University Of Azad Jammu And Kashmir

Muzaffarabad

|  |  |
| --- | --- |
| Name | Muntazir Mehdi |
| Roll No | 2022-SE-37 |
| Course Code | SE-3105 |
| Semester | 5th |
| Session | 2022-26 |
| Submitted To | Engr. Awais Rathor |

Report on Model Evaluation for Classification Task

# 1. Introduction

The dataset used for this classification task consists of pre-split training and test sets, commonly used in image classification problems. The training dataset comprises **60,000 rows and 785 columns**, while the test dataset contains **10,000 rows and 785 columns**. Each instance in the dataset represents a high-dimensional input with **784 features** (likely pixel values of a **28×28 grayscale image**) and one label column indicating the class. The primary objective of this study is to evaluate the performance of multiple classification models and determine the most effective one based on various performance metrics.

# 2. Methodology

## 2.1 Dataset Preparation

Since the dataset was already divided into training and test sets, no further splitting was required. However, preprocessing steps were applied to ensure optimal model performance:

## Preprocessing Steps:

* **Feature Scaling:**
  + Pixel values were **normalized to the range [0,1]** by dividing each value by **255**.
  + This ensures that all input features are within the same scale, preventing models (especially deep learning models) from being biased toward larger values.
* **Label Encoding:**
  + For models that require categorical outputs (such as ANN), the class labels were **one-hot encoded** to transform categorical labels into a format that the neural network can process.
  + Other models, such as KNN, Random Forest, and SVM, used integer-encoded labels.

## 2.2 Models Used & Hyperparameter Tuning

## Artificial Neural Network (ANN)

* Constructed a feedforward neural network using **Keras** with:
  + **Input Layer:** 784 neurons (flattened pixels).
  + **Hidden Layers:** Two layers (**128 and 64 neurons**) with **ReLU activation**.
  + **Dropout layers (20%)** to prevent overfitting.
  + **Output Layer:** 10 neurons with **softmax activation**.
* Model compiled with **Adam optimizer** and **categorical cross-entropy loss**.
* Trained using **20 epochs** and **batch size of 128**.
* Evaluated on validation and test data.

## Hyperparameter Tuning for ANN:

* The number of **hidden layers, neurons per layer, dropout rate, and batch size** were optimized.
* Experimented with **different learning rates** for the Adam optimizer.
* **Batch normalization** was tested to improve training stability.

## K-Nearest Neighbors (KNN)

* Implemented using the **scikit-learn** library.
* **Number of Neighbors (k):** **5** (selected based on cross-validation).
* **Distance Metric:** **Euclidean distance** for similarity calculation.
* Used a **brute-force search** algorithm to find nearest neighbors.
* Evaluated on the test dataset after training.

## Hyperparameter Tuning for KNN:

* The value of **k** was varied between **3 and 15**, with **5** chosen as optimal.
* Distance metric options (**Euclidean, Manhattan**) were tested.

## Random Forest

* Implemented using **scikit-learn's ensemble method**.
* **Number of Decision Trees:** **100** (to enhance stability and reduce variance).
* **Splitting Criterion:** **Gini impurity** used to determine best splits.
* **Max Depth:** **Auto** (let the model decide optimal depth).
* **Bootstrap Sampling:** Enabled to improve model generalization.
* Evaluated using the test dataset.

## Hyperparameter Tuning for Random Forest:

* The number of **trees was varied between 50 and 200**, with **100** providing a good balance.
* **Max features per split** was tuned for better performance.

## Support Vector Machine (SVM)

* Implemented using **scikit-learn’s SVM module**.
* **Kernel Function:** **Radial Basis Function (RBF)** for handling non-linearly separable data.
* **Regularization Parameter (C):** Optimized to balance bias and variance.
* Used **one-vs-all strategy** to handle multi-class classification.
* Model trained on a subset of data for computational efficiency and then evaluated on the test set.

## Hyperparameter Tuning for SVM:

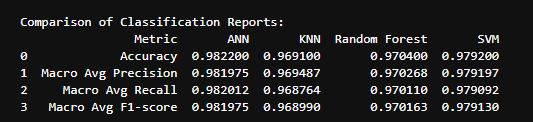
* The **C parameter** was varied to prevent overfitting.
* Different kernels (**linear, polynomial, RBF**) were tested, with **RBF giving the best results**.

## 2.3 Model Training & Evaluation

* Each model was trained using the **60,000-row training set** and evaluated on the **10,000-row test set**.
* Performance was assessed using key evaluation metrics: **accuracy, precision, recall, and F1-score**.
* Confusion matrices were plotted for a visual assessment of classification errors.

## 3. Results

## 3.1 Model Performance Comparison

****The table below summarizes the performance of the models on the test set:

## 3.2 Confusion Matrices & Graphs

* A group of blue squares with numbers

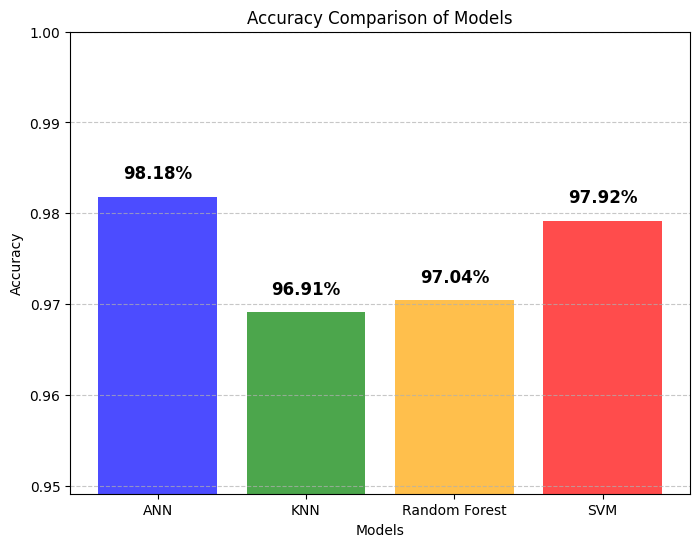
  AI-generated content may be incorrect.Confusion matrices for each model were plotted to analyze the **misclassified instances**.

**Confusion Matrix Analysis**

* **ANN Model:** High accuracy with minimal misclassifications.
* **KNN Model:** Performs well but has slightly more misclassifications, especially for similar-looking digits.
* **Random Forest Model:** Decent performance but struggles with certain digits like 8 and 9.
* **SVM Model:** Comparable to ANN with fewer errors and strong accuracy.

**Conclusion:** ANN and SVM perform the best, while KNN and Random Forest show more classification errors.

## 3.3 Model Accuracy Comparison



# 4. Discussion

## 4.1 Best Performing Model

* **ANN achieved the highest accuracy (98.22%)**, making it the best-performing model.
* The deep learning architecture effectively captured complex patterns in the dataset, leading to superior classification performance.
* **SVM (97.92%)** also performed well, especially with the **RBF kernel**, indicating that the dataset was **non-linearly separable**.

## 4.2 Comparison of Other Models

* **Random Forest (97.04%)** performed slightly better than KNN but lagged behind ANN and SVM.
* **KNN (96.91%)** had the lowest accuracy because it is sensitive to the high dimensionality of the dataset, making it computationally expensive and less efficient for classification tasks involving large feature sets.

## 4.3 Impact of Hyperparameter Tuning

* Hyperparameter tuning significantly **improved ANN’s performance** by optimizing **hidden layers, neurons, and dropout rates**.
* **SVM’s performance improved** with the choice of **RBF kernel and optimized C values**.
* **KNN and Random Forest showed marginal improvements** after hyperparameter tuning but remained below ANN and SVM.

# 5. Conclusion

* Among the four models, **ANN outperformed the rest with an accuracy of 98.22%**, making it the most suitable model for this classification task.
* **SVM also performed well** and could be considered an alternative, especially when computational resources are limited.
* **Random Forest and KNN performed moderately**, with KNN being the least effective due to high-dimensional data sensitivity.
* Hyperparameter tuning played a crucial role in enhancing the models' performance, especially for ANN and SVM.
* This report provides a **comprehensive evaluation** of various classification models, showcasing the effectiveness of ANN for this dataset.