Machine Learning Classification Report

# Introduction

The objective of this assignment is to apply various machine learning models and techniques to classify numerical and categorical data. This task is crucial in the field of machine learning as it demonstrates the application of different algorithms and preprocessing techniques to achieve optimal performance.

# Dataset Overview and Preprocessing

The datasets used for this task include both numerical and categorical data. Preprocessing steps involved handling missing values, converting categorical features to numerical features, and splitting the data into training and testing sets.

## Data Preprocessing

# Data Preprocessing

# Implementation of Models

The team has implemented the following models:  
(a) Logistic Regression, Decision Tree, Random Forest, and Artificial Neural Network for the classification problem.  
(b) A Convolutional Neural Network for the digit recognition problem.

# Data Preprocessing

The team has performed data preprocessing tasks such as loading the data, handling missing values, and converting categorical features to numerical features.

# Model Fine-tuning

The team has tuned the hyperparameters of the models to improve performance using techniques such as GridSearchCV.

# Data Visualisation

The team has visualised the data using appropriate plots to understand the features and the target variable.

# Performance Evaluation

The team has evaluated the performance of the models using appropriate metrics such as accuracy, confusion matrix, and cross-validation scores, and reported the results.

# Collaboration

The students have effectively collaborated on the assignment, with each student's contributions clearly documented.

# Model Exploration

The team have explored different model architectures or hyperparameters beyond the basic requirements.

# Introduction

The objective of this assignment is to apply various machine learning models and techniques to classify numerical and categorical data. This task demonstrates the application of different algorithms and preprocessing techniques to achieve optimal performance.

# Data Preprocessing

The data preprocessing involved handling missing values, converting categorical features to numerical features, and splitting the data into training and testing sets. The specific steps included:

- Loading the data

- Handling missing values

- Converting categorical features to numerical features using techniques like one-hot encoding

- Splitting the data into training and testing sets

# Implementation of Models

The team implemented several models for classification:

- Logistic Regression

- Decision Tree

- Random Forest

- Artificial Neural Network

Additionally, a Convolutional Neural Network (CNN) was implemented for a digit recognition problem.

# Model Fine-Tuning

Hyperparameters of the models were tuned to improve performance using techniques such as GridSearchCV. For example, for the Random Forest model, the best parameters found were `n\_estimators=100` and `max\_depth=10`.

# Data Visualization

Data visualization was performed using appropriate plots to understand the features and the target variable. This included:

- Histogram plots

- Scatter plots

- Box plots

# Performance Evaluation

The models' performance was evaluated using metrics such as accuracy, confusion matrix, and cross-validation scores. For example, the accuracy of the Random Forest classifier was 96.89%, and the confusion matrix and cross-validation scores were as follows:

- Confusion Matrix:  
```  
 Predicted 0 Predicted 1  
Actual 0 128 8  
Actual 1 7 135  
```

- Cross-Validation Scores:  
```  
Cross Validation Scores: [0.97837838 0.96756757 0.95135135 0.95675676 0.91351351]  
Average CV Score: 0.9535135135135135  
```

# Model Exploration

The team explored different model architectures and hyperparameters beyond the basic requirements to optimize performance further.

# Collaboration

The students effectively collaborated on the assignment, with each student's contributions clearly documented.