**Assignment No. 3**

|  |
| --- |
|  |

**Part 1**

Adjust <https://github.com/rasbt/machine-learning-book/blob/main/ch11/ch11.ipynb> to use two layers for classifying handwritten digits MNIST dataset using the same full ANN architecture presented in the class and evaluate its prediction performance (macro AUC) using Train/Test validation procedure.

**Evaluation**

In this section, we compare the performance of the above implementations (Bagging Classifier) with the implementation in sklearn package (DecisionTreeClassifier). In order to evaluate the performance of the methods, we used classification datasets available on the internet (www.kaggle.com).

1. Water Quality Dataset (<https://www.kaggle.com/datasets/mssmartypants/water-quality>)
2. Titanic Dataset (<https://www.kaggle.com/competitions/titanic/>)
3. Heart Failure Prediction Dataset (<https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction>)
4. Cancer Dataset (<https://www.kaggle.com/datasets/erdemtaha/cancer-data>)
5. Diabetes Binary Classification (<https://www.kaggle.com/datasets/girishvutukuri/diabetes-binary-classification>)

**Hyper Parameters** – We chose to…

1. n\_estimators\_n = 300
2. max\_samples\_n = 1
3. max\_features\_n = 0.5
4. depth\_n = 5

**Preprocessing** – In order to enable the use of those methods on the abovementioned datasets, preprocessing was needed. This include for each dataset, dropping incomplete records, binarize non-numerical columns, and converting the X and y using existing function in sklearn, *lsklearn.preprocessing.KBinsDiscretizer*, an algorithm that implements discretization (also known as quantization or binning) to transform continuous features into discrete ones.

**Evaluation Metrics** – We compared the performance of each model using evaluation metrics by using repeated K-fold cross-validation to evaluate both algorithms. We evaluated each model using the following performance measures:

1. Accuracy
2. Precision
3. Recall
4. F1-score
5. ROC-AUC score

Data analysis

Write a short report presenting the detailed results and summarizing your findings, including which model performed the best, and why.

The report should include a table that compares the predictive performance of the various methods, in the following structure:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Dataset** | **Method** | **Evaluation Metric** | **Evaluation Value** | **Fit Runtime (ms)** |
| Water Quality | BaggingClassifier | Accuracy |  |  |
| Precision |  |  |
| Recall |  |  |
| F1-score |  |  |
| ROC-AUC score |  |  |
| DecisionTreeClassifier | Accuracy |  |  |
| Precision |  |  |
| Recall |  |  |
| F1-score |  |  |
| ROC-AUC score |  |  |
| Titanic | BaggingClassifier | Accuracy |  |  |
| Precision |  |  |
| Recall |  |  |
| F1-score |  |  |
| ROC-AUC score |  |  |
| DecisionTreeClassifier | Accuracy |  |  |
| Precision |  |  |
| Recall |  |  |
| F1-score |  |  |
| ROC-AUC score |  |  |
| Diabetes Prediction | BaggingClassifier | Accuracy |  |  |
| Precision |  |  |
| Recall |  |  |
| F1-score |  |  |
| ROC-AUC score |  |  |
| DecisionTreeClassifier | Accuracy |  |  |
| Precision |  |  |
| Recall |  |  |
| F1-score |  |  |
| ROC-AUC score |  |  |
| Heart Failure Prediction | BaggingClassifier | Accuracy |  |  |
| Precision |  |  |
| Recall |  |  |
| F1-score |  |  |
| ROC-AUC score |  |  |
| DecisionTreeClassifier | Accuracy |  |  |
| Precision |  |  |
| Recall |  |  |
| F1-score |  |  |
| ROC-AUC score |  |  |
| Airline Passenger Satisfaction | BaggingClassifier | Accuracy |  |  |
| Precision |  |  |
| Recall |  |  |
| F1-score |  |  |
| ROC-AUC score |  |  |
| DecisionTreeClassifier | Accuracy |  |  |
| Precision |  |  |
| Recall |  |  |
| F1-score |  |  |
| ROC-AUC score |  |  |