

Machine Learning

Digits Data Set – Perceptron Classifier

Perceptron Classifier

Use the digits dataset available under SKLearn.

Consider the data corresponding to classes 0 and 1 only.

Each pattern is a 8 X 8 sized character where each value is an integer in the range 0 to 16.

Convert it into a binary form by replacing a value below 8 by 0 and other values (≥ 8) by 1.

Split the dataset into train and test parts.

Do this splitting randomly 10 times and report the average accuracy.

CODE:

Please find the code committed for Perceptron Classifier

[*Perceptron_DigitsDataSet_Impl.py*](#)

- Perceptron classifier from sklearn linear model is used to classify the data in this subtask.
- The average accuracy is computed for 10 trials.

RESULT:

Average Accuracy for Perceptron Classifier for test_size= 0.1 : 0.9944444444444442
Average Accuracy for Perceptron Classifier for test_size= 0.2 : 0.9972222222222221
Average Accuracy for Perceptron Classifier for test_size= 0.3 : 0.9958333333333333
Average Accuracy for Perceptron Classifier for test_size= 0.4 : 0.9944444444444445
Average Accuracy for Perceptron Classifier for test_size= 0.5 : 0.9930555555555556

INFERENCE/ANALYSIS:

- **Perceptron classifier** is a linear algorithm that can be applied to binary classification tasks.
- Learning rate is one of the important hyperparameters in the perceptron model and can decide whether the model can learn fast at the cost of lower skill or learn slow for better results.
- **The overall accuracy is high for different test sizes as can be seen from the test result as this is a binary classification model with a relatively smaller dataset.**

RESOURCES USED FOR THE ASSIGNMENT:

- **Environment:**
Anaconda, Jupyter notebook
- **Software:**
Python
Python libraries/modules: Pandas, Numpy, SkLearn, etc.