**NN & DL - IP5**

**Revanth Bharadwaj Tarimela Roopavataram**

**Student ID – 700743276**

**Github Link:** [**https://github.com/revvanthhh/ICP5**](https://github.com/revvanthhh/ICP5)

1. Implement the Naïve Bayes method using the scikit-learn library

* Use dataset available with name glass.
* Use train\_test\_split to create the training and testing part.
* Evaluate the model on test part using score and classification\_report(y\_true, y\_pred)

A screenshot of a computer

Description automatically generated

Text

Description automatically generated

This example divides the glass dataset into feature variables (X) and target variables after loading it into a pandas data frame (y). The train test split function is then used to divide the data into training and testing sets. The training data are used to train the Gaussian Naive Bayes model, which is subsequently applied to the test data to create predictions. The model is assessed using the classification report and accuracy score.

2. Implement linear SVM method using scikit library

* Use the same dataset above.
* Use train\_test\_split to create the training and testing part.
* Evaluate the model on the test part using score and classification\_report(y\_true, y\_pred)

Graphical user interface, text

Description automatically generated

Text

Description automatically generated

This example divides the glass dataset into feature variables (X) and target variables after loading it into a pandas data frame (y). The train test split function is then used to divide the data into training and testing sets. The training data are used to train the linear SVM model, which is then used to the test data to provide predictions. The model is assessed using the classification report and accuracy score.

3. Which algorithm did you get better accuracy with? Can you justify why?

Ans. Naive Bayes is a simple and fast algorithm that makes strong assumptions about the independence of features, which can be a good option for high-dimensional datasets with a lot of features. On the other hand, linear SVM is a more complex algorithm that can handle non-linear decision boundaries, which can be a good option for datasets with a complex relationship between features and the target variable. Based on the accuracy, the best algorithm for the above problem is SVC algorithm. We can consider this by checking the factors in the output.