

# INFO

---

## **Libraries:**

The external libraries used in Assignment 2 include scikit-learn, nltk.

For each method used in the code, comments have been put which depict what the function used does.

Here are some of the important methods used:

1) `MLPClassifier()`: It is the implementation of neural networks in scikit-learn module. The hidden layers size is 100 by default. The learning rate is 0.001 by default.

2) `word_tokenize()`: It's a part of nltk which tokenizes the line of strings into words.

3) `Predict()`: This is used for testing the classifier we have developed.

4) `metrics.accuracy_score(predicted, groundtruth)`: This gives us the Accuracy of the testing dataset that predicts using a particular classifier.

## **Challenging Parts:**

The challenging part of the assignment is the preprocessing of the String data in Amazon Review dataset and to improve the execution time of the algorithm for the amazon dataset. It takes almost 2 hours to finish the execution.

## **Easy Parts:**

The easy part is Digit Recognition dataset. This was one of the easy part used to predict the labels.