Habib University
CSE 351 - Artificial Intelligence
Fall 2020
Assignment 2 – Part A

Objective:

This assignment gives students hands-on experience with learning and classification a classifier. The assignment will also familiarize them with some machine learning libraries in python.

Q1 – Classification [45 points]

Problem 1.1 – Implementing Naïve Bayes [35 points]

You are given a dataset of textual summary of medical queries classified into five different categories. You have to build your own naïve Bayes classifier to predict these categories for future queries. The task includes:

- **[15 points]** Implement Learn(..) method that takes training and test data (in the form of word vectors) and learns its Naïve Bayes classifier. You have to do your own implementation of Naïve Bayes here.
- **[05 points]** Some conditional probabilities may turn out to be zero in the training dataset. Modify your classifier such that it applies Laplacian smoothing while learning conditional probabilities to avoid such zero values.
- **[05 points]** Implement Predict(…) that takes a learned classifier and test data and returns predicted values.
- **[10 points]** Implement Evaluate(…) that takes actual and predicted labels and returns precision, recall, f-measure. The method will also display confusion matrix.

The code will be written in the attached Ass2_classification.ipynb. You can find here details of NaïveBayes classification.

Problem 1.2 - [10 Points]

You are given a Fashion-MNIST dataset comprising of 28x28 grayscale images of 60,000 fashion products from 10 categories, with 7,000 images per category. The training set has 60,000 images and the test set has 10,000 images. Some sample images are shown below:

The dataset is freely available here.

You have to build a scikit-learn based classifier for this dataset to classify each input image into one (out of 10) categories. You are required to experiment with three to four different classifiers and report accuracy of each of them.

Note: Your code will follow the same structure given in Problem 1.1.

Submission Instructions

Submissions will be made on the LMS by the due date (announced on LMS). No email submission will be accepted. The submitted file should be in the form of a ZIP file named as **<studentid>_Ass2** containing separate folders named Q1 and Q2 for the source code. The source code should be written in the attached notebooks.