

Habib University
CSE 351 - Artificial Intelligence
Fall 2022
Assignment 2 – Supervised Learning

Objective:

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

Q1 – Learning an Artificial Neural Network [30 points]

You are given a dataset that represents KSE Index values in past few years. You are required to learn a neural network that predicts the value of KSE Index based on its past three values. There are four columns in the data, namely, Index(t-3), Index(t-2), Index(t-1) and Index(t), represented as X1, X2, X3 and Y, respectively. Your job is to learn Index (t) as a function of Index(t-1), Index(t-2) and Index(t-3). In other words, your job is to predict the value of tomorrow's index value using the values of today, yesterday and day before yesterday. Your task is divided into the following steps:

- **[05 points]** Load dataset from the CSV file and splits it into train (67%) and test (33%) parts.
- **[15 points]** Implement Learn(..) method that will learn and return a neural network classifier. You have to do your own implementation of ANN here. The method will take the following parameters:
 - o training data
 - o number of input, hidden and output neurons
 - o learning rate
 - o number of epochs
- After every epoch, the neural network will compute Mean Squared Error (MSE) over the test data and will print it.
- **[05 points]** Implement Predict(...) that takes a learned classifier and test data and returns predicted values.
- **[05 points]** Implement computeError(...) that takes actual and predicted values and return Mean Squared Error.

The code will be written in the attached Ass2_Q1.ipynb.

Q2 – Text Classification using Naïve Bayes [30 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:

- **[10 points]** Implement Learn(..) method that takes training (in the form of word vectors) and learns a Naïve Bayes classifier (i.e. all probabilities that a Naïve Bayes classifier needs). 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_Q2.ipynb. In addition to the lecture slides of Naïve Bayes classifier, you can also refer to [this link](#) for further details on using Naïve Bayes for text classification.

Q3 – Image Classification in Fashion Retail [15 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 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.

The code will be written in the attached Ass2_Q3.ipynb.

Submission Instructions

The assignment will be submitted on 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 <studentid1>_<studentid2>_Ass2 containing separate folders for each question.