Artifical Intelligence

Assignment No.3

Due Date/Time	27th May, 2023 11:30 PM
Files to be submitted	Documentation along with code file.
File Naming	Roll_No_Assign_03.pdf Roll_No_Assign_03.ipynb Note: Any assignment that is not according to format will be marked as zero.
Coding Guides	**1.** Use of proper variable declaration/initialization according to the naming conventions (**camelCase, snake_case, PascalCase**) **2.** Use of proper function for each module. Note: Marks will be deducted if not following the above guide line.
Submission Guide line	**1.** Code along with documentations should be submitted on teams by due date/time.
Plagiarism	Any kind of plagiarism will result in F grade in course
Weightage	This assignment will be marked on CLOs, also it will be graded for course and lab both. Which means , if you have got 8/10 in this assignment same marks will be reflected for your lab assignment.

***Group of two members are allowed for this assignment**

The goal in this assignment is to develop a machine learning model that can classify between images of T-shirts and dress-shirts.

You are given the following files:

• **TrainData.csv**: It contains 12000 training examples. Each row contains 784 values. The dataset has been derived from Fashion-MNIST dataset. Each example is a flattened 28x28 pixel gray-scale image. You can reshape the examples to visualize what each image looks like.

• **TrainLabels.csv:** This file contains true labels for the examples in TrainExamples.csv

• **TestData.csv:** This file contains test examples.

• **You can load train and test data using the following code:**

```
Xtr=np.loadtxt("TrainData.csv")

Ytr=np.loadtxt("TrainLabels.csv")

Xts=np.loadtxt("TestData.csv")
```

• To visualize an example (say trainining example at index 10, you can use the following code):

```
import matplotlib.pyplot as plt

plt.imshow(Xtr[10].reshape([28,28]))
```

Tasks:

- **Write a method named ExtractFeatures.**

It should take as arguments the examples, compute and return a feature vector, that you think would help in performing the task, corresponding to each example. **You have to create your own two function at least for extracting the feature they can be edges,color channels,or midpoints (Apply simple filtering techniques that you have studied in CV to extract features from an image)**

- You are supposed to train a model using your extracted features. **Try at least 2 different classification techniques of your choice.**

- **Using 5-fold cross-validation**, optimize hyperparameters for the models. Since the dataset is balanced, you can use classification accuracy as the performance metric.

- **After choosing the best hyperparameters**, use the complete training dataset to train the final model. Dump the model in the file named “myModel.pkl”. You can take help from https://scikit-learn.org/stable/modules/model_persistence.html for saving your model.

- **Create a file test.py**. It should read the test examples from **TestExamples.csv**. DO NOT CHANGE THE ORDER OF EXAMPLES. Generate predictions for the test examples and save them to a csv file named “myPredictions.csv”. You can use the following code to save your results: np.savetxt("myPredictions.csv", Yts)

You have to submit the following files:

- Files containing the **cross-validation code** used for hyperparameter selection.

- **Training.py:** The file containing the code to train and dump the final model using the chosen hyperparameters.

- **Testing.py:** The file in which you load the saved model and generate and save predictions for test data.

- **myPredictions.csv:** The csv file containing your predicted labels for the test set.

A pdf report that contains the following information:

- What features did you use and why?
- What classification techniques did you try?
- Which of the methods (and for what hyperparameters) showed best cross validation performance.
- What test accuracy are you expecting?

Note: You may use any libraries or modules in Python that you find helpful, but you should be able to explain how they work and why you used them in your code