```
Data Science: A Programming Approach
Mahyar S Vaghefi
University of Texas Arlington
```

Individual Project - Fall 2020

You need to work on a popular Fashion MNIST dataset for this project. The dataset includes tiny images of fashion pieces. The objective is to create a set of supervised learning models that can predict the type of item based on its image. You can use all different models that you learned about them in this course for yourr work. Keep in mind that this is a project, not a class assignmnet. So, not all steps are predetermined and you have more flexibility, and the final outcome is likely to be more detailed.

In order to load the dataset you need to have `tensorflow V2` on your computer. Use the following code to install the package

```
In [ ]: # !pip install --upgrade tensorflow
```

You can also check the version of it using the following code.

```
In [18]: import tensorflow as tf
         tf.__version__

Out[18]: '2.1.0'
```

Now, it's time to load the dataset

```
In [19]: from tensorflow import keras
         fashion_mnist = keras.datasets.fashion_mnist
         (X_train, y_train), (X_test, y_test) = fashion_mnist.load_data()
```

As can be seen from the above code, the dataset was divided into train and test sets. Let's take a look at the X_train

```
In [20]: X_train.shape

Out[20]: (60000, 28, 28)
```

As it is clear, the train dataset (`X_train`) contains 60,000 images of size 28 x 28. We can visualize one of the images using the following code:

In [22]:

```
import matplotlib as mpl
import matplotlib.pyplot as plt
%matplotlib inline

sample_image = X_train[1]
plt.imshow(sample_image, cmap='binary')
plt.axis('off')
plt.show()
```

The `y_train` also includes values between 0 and 9. Each represents a particular category. For example, we can check the value of `y_train` for the above image.

In [25]:

```
y_train[1]
```

Out[25]:

```
0
```

The above code shows that the image belongs to category 0. To get the associated label with each category, you can use the following code:

In [26]:

```
class_names = ['T-shirt/top','Trouser','Pullover','Dress','Coat','Sandal
print(class_names[y_train[1]])
```

```
T-shirt/top
```

Now, it's your turn,

- **Task1:** Use the train set to train various supervised models and evaluate their performance using the test set.
  - Use different supervised learning models.
  - Use different metrics such as **accutacy**, **precision**, **AUC**, and ... in your model evaluation.
  - It is not enough to report the metrics. It is crucial that you interpret the metrics for each model and compare them across different models.
  - You may need to use the cross validation methods for hyperparameter selection.
  - Specify the model that outperforms the other models.
- **Task2:** Use the best model to predict your own fashion pieces.
  - Take a picture of five fashion pieces of your own (take pictures in square format).
  - Resize images to the correct size (28,28).
  - Grayscale your images.
  - Visualize all the images side by side
  - Use the best model in Task 1 to predict the label of each of your own images.
  - How accurate is the final result?

Output

- Make sure to put descriptive comments on your code
- Use the markdown cell format in Jupiter to add your own interpretation to the result in each section.
- Make sure to keep the output of your runs when you want to save the final version of the file.
- The final work should be very well structured and should have a consistent flow of analysis.

Due Date: Nov 9 2020 at 11:59 PM

Grading Criteria

Comprehensiveness	**30%**
Correctness	**20%**
Complete Report	**20%**
Clear Code	**20%**
Innovation (Extra)	**20%**
Total	**110%**