Image Classification with Supervised Machine Learning Techniques

Assignment Scenario:

Mr. Joel recently met a girl on Tinder named Suzy Bae. However due to COVID-19 restrictions and social distancing, he only video calls and shares pictures with her. This is his first relationship and he is unsure about many things. He is very unsure whether she is having a bad day or a good day through her pictures and video calls. He seeks your help as AI experts to help solve this problem. Your group has the big responsibility to write a program that can help classify her pictures that she sends him.

Figure 1: Joel and Suzy

Introduction

In this assignment you will go through the different stages of machine learning which includes data preparation, creating and training a classifier, model evaluation as well as optimisation. The assignment shall involve creating a Python image classification program that can:

- Accept image datasets as input
- Prepare and pre-process the datasets
- Implement various machine learning algorithms (classifiers)
- Train and test classifiers on the input datasets
- Evaluate models produced from the classifiers

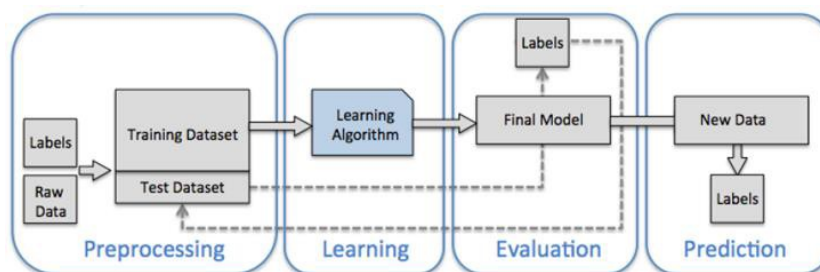To achieve this, you would first need to understand the machine learning workflow:

Figure 1: Machine Learning workflow

1. ***Data preparation:*** A typical machine learning project starts with data preparation. During this stage, a dataset containing images and class labels is compiled. For example, a dataset can be as simple as having two folders, one named cat and another named dog, each filled with images of cats and dogs downloaded from the Internet.
2. ***Data pre-processing:*** Once a dataset has been compiled, it would need to be pre-processed. For instance, the downloaded cats and dogs images may be of varying sizes, therefore, they would need to be cropped such that they are of the same dimensions. In addition to that, the dataset

3. **_Training:_** The next step is to apply a learning algorithm on the dataset so that significant patterns (features of cats and dogs) are learnt from the training set. Since the images and labels are provided, a supervised learning algorithm should be implemented.
4. **_Evaluation:_** At the end of training, a model would be produced by the learning algorithm. This model needs to be evaluated by performing predictions on the unseen testing set. Images of cats and dogs would be passed to the model, and based on how many it guessed correctly, the model can be deemed accurate or inaccurate.

Program

- You should code your program in Python (version 3+)
- These are some Python libraries that will be helpful:
  - scikit-learn: https://scikit-learn.org/stable/
  - Matplotlib: https://matplotlib.org/
- Your program should have these modules:
  - Data input – prepares and pre-processes datasets
    - You need to download and run using at least the **first two** datasets for training, you can also test and train just on the selected database for initial results
      - FEI Face Database (frontal at least): https://fei.edu.br/~cet/facedatabase.html
      - JAFFE Database: https://zenodo.org/record/3451524#.XrQQUqgzY2x
      - Your own choice of dataset [Extra challenge]
        - Can be downloaded from a website
        - Or compiled by yourself
    - You need to set your own labels or separate the data into different folders
    - Try out different combinations or partitions of data for training for instance at least:
      - A) FEI alone for training
      - B) JAFFEE alone for training
      - C) FEI + JAFFE

  - Machine learning/Deep Learning – creates and trains classifiers
    - Your program should implement at least **two algorithms (ML & DL)**
  - Model evaluation – evaluates models generated from classifiers
    - Testing data (Joel's private collection), you also can add more data or make some changes (please justify)
  - https://drive.google.com/open?id=1Yre-YydolCOQuNxU8P_P4kEWq0NNdqjD

    - Your program should implement at least **two methods of evaluation**

* For the statements, opinions, and suggestions that you make above, please provide a reasoned argument for it. Do not just state something and assume that it is true. Provide explanation and justification for the statements that you make.
1.