Assessment 2

This assessment concerns a very standard task in machine learning: image recognition. We have done image recognition in-class so you should familiar with the general problem. The problem given to you here is slightly more challenging than the one in-class. Hence, you are in groups of 2.

Please very carefully read through this document as not doing so may result in significant loss of grade and even a fail.

Submission Time & Date: 11:00 pm on 29th September, 2019

Section I: Defining the Problem

Context

You are still part of the same organization as in Assessment 1. Now you are working in a team with one other machine learning expert.

Problem

Your scientist colleagues have taken millions of photos of very small particles as part of their experiments. The size of the images is 32x32. In these images, your colleagues expect to see 10 different particles. Your colleagues want your team to build a system which automatically recognizes the particles in each image and labels the image with that particle. Assume that there is a just one particle in each image.

As preparation for this, you take a labeled dataset 32x32 images. Each image contains just one object of interest and there are a total of 10 objects. The details of the dataset are given later. You will first build a system for these images as a test run for the images given by the scientists.

Task

Your task is to design and train a neural network which will accurately identify the objects in the images in the given dataset. The architecture and optimization of the neural network is completely your decision. However, the neural network should take a 32x32 image in the format identified in the dataset section, and in the output identify the object.

Dataset

The dataset is contained in the images.zip folder. The data is split among 5 batch files with each batch file containing 10000 images. There is also a meta file which contains labeling information, i.e. which label refers to which object. You can import the files in python using the following code:

```
import pickle
with open(file, 'rb') as fo:
    dict = pickle.load(fo, encoding='bytes')
```

You can read each file by using the appropriate filename. Running this code will return a dictionary containing the data. The dictionary will have four keys with two keys referring to the input pixels and the output labels. The rest of the data exploration is left upto you.

Section II: Assignment Submission

Required Files

You are required to submit three files. The three files must exactly conform to the requirements below otherwise you will lose grade. Unfortunately, in assessment 1, many students did not follow the guideline. We were very accommodating in assessment 1. This time we will not be so accommodating. **If the files do not follow the requirements, we will not be able to mark them and award an F**. The required files are

1. A neural network model file (.h5) which contains your neural network model. The model file must be named 'ourmodel.h5' (quotation marks not included).
2. A python script (.py file) which contains all your code work and shows the process through which you developed your solution. The code should be broken down into sections and with appropriate comments making it easy to follow. You will lose grade if your code is not easy to follow. The source code file should be named 'source_code.py'.
3. A report summarizing your design and detailing how you arrived at your solution, what difficulties you faced and how did you try to tackle those difficulties. Your report should follow the template provided.

Remember: You will receive a Fail grade by default if you don't submit all three files listed above. You will also receive a Fail grade if the files are not submitted according to the specifications listed below. It is your responsibility to make sure that the files are submitted correctly. Make sure you submit the files well before deadline.

File Submission

Your files should be submitted exactly as given below otherwise **you will receive a Fail**:

1. The report should be submitted through the Turnitin link. It should be of pdf format. The pdf file should be named 'first student last name_second student last name.pdf'(quotations marks not included). For example, if I submitted the file it would be named Zafar_123456.pdf
2. The source code file as well as your model file should be submitted together in a zip folder in the Assignment Submission link. The folder name should be the same as document name (of course the file type will be different).

Section III: Our Expectations

There are many image recognition solutions available on the web. You are welcome and encouraged to explore them. However, we expect you to make your own solution and spend considerable time designing it to increase it is accuracy. We will be able to tell from your report and code how much effort you put into the solution. We will also be able to judge your effort and knowledge in the oral exam in assessment 3. We expect that you have learned from your mistakes from assessment.

This is a challenging assessment so please start on it as soon as possible. Don't leave it till last week. Ideally you should be spending 3 - 4 hours every week and should finish your assessment 2 -3 before deadline. If you are working every week and seriously, we expect you to run into challenges, and even after considerable effort form your side you may not able to solve them. We expect you to come to

us for guidance then. But we can't/won't help you if you have not done something yourself or are coming to us in the last week.

Section IV: Marking Criteria

The marking of this assessment will go hand-in-hand with marking of assessment 3. Assessment 3 is an oral exam relating mostly to assessment 2. If you cannot answer questions about your solutions in assessment 3, it will inform us that you did not solve the problem yourself and we will adjust your assessment 2 mark accordingly.

If you meet the minimal requirements for submission above, you can achieve the following grades

Pass: To achieve a pass mark you must show a basic understanding of the process and tasks involved in solving the given problem and have a basic working solution which gives good accuracy and has a decent run time. Furthermore, your source code should be well commented and broken down into sections. Lastly, your report should be well-written.

Credit: To achieve a credit you must show a good understanding of the process and the tasks involved, and have a working solution with a superior accuracy and a good run time. Moreover, you can partially identify challenges of the problem, the implementation issues, and some difficulties you faced and how to tackle them.

Distinction: To achieve a distinction you must show an excellent understanding of the process, the tasks involved, and have a well-implemented working solution with excellent accuracy and excellent run-time. Furthermore, you can identify the majority of the interesting features of the problem, the implementation issues the requirements pose and how you solved all these issues.

High Distinction: Everything in distinction but at an outstanding level and even going beyond!