Machine Learning Online Assignment

Question-1 Working with Pokemon Dataset | Learning to use Kaggle API

1. Download the Pokemon Dataset from Kaggle (https://www.kaggle.com/kvpratama/pokemon-images-dataset). Learn how to work with Kaggle API. Don't use the Download Button, download via command line as it would be useful when you download datasets on cloud servers later for training purposes. **[10 points]**

2. Visualise and randomly shuffle this dataset. Draw any 10 unique pokemons. **[10 points]**
   We are not going to do any classification, since many pokemons have only 1 image. Howerver, feel free to try train any SVM/Logistic classifier.

Question-2 One vs Rest Scheme Implementation | CIFAR-10 Dataset

1. Download the CIFAR-10 dataset(Use only first batch containing 10,000 images or use all if you have computing power)from this link - https://www.cs.toronto.edu/~kriz/cifar.html Draw a bar-chart showing frequency count of each of the 10 classes. **[10 points]**

2. What are the dimensions of the data loaded? Visualise the images, you should be able see car, truck, animal etc. (Hint- You might need to change the axis to have H X W X Depth format)**[20 points]**

3. We have learned that SVM is a Binary Classifier, but the current problem is an example of Multiclass Classification - and we use generally use **One Vs Rest** or One vs One schemes to do such a classification. Now you have already seen the **One vs One Scheme** implementation in the lectures,
   Define a SVM classifier and train your model using **One vs One Rest** scheme to make predictions on the dataset. Tune your learning rate, plot the loss for any two classes as well. You can use the Pegasos implemenation which implements a Linear SVM. **[50 points]**

5. Repeat the above part with Sci-kit learn SVM and compare your results. Draw the confusion matrix using sk-learn confusion. Try experiementation with different kernels, Use Grid Search over various values of C and Kernel Type. **[20 points]**

6. Generate classification report using sk-learn classification report function. Refer

https://scikit-learn.org/stable/modules/generated/sklearn.metrics.classification_report.html
[10 points]

For any clarifications, please post on Discuss