Theory and Deep Learning Coding HW 1 (Report)

Steps to run code

Before you continue, check that your python environments contains the following dependencies:

- numpy
- sklearn
- pandas

There are two alternatives:

1. If you prefer Jupyter notebook:
    a. Run jupyter notebook on your terminal where the directory is the root directory of the “Coding_HW1_2020_Nix” folder
    b. There are two notebooks. “DL1_Splitting.ipynb” contains the code for splitting the dataset, where the .npy files of the split data is stored in the “train_val_test_sets” folder. “DL1_Training” contains the code for the training and testing.
    c. The notebooks are documented for you to follow through.
2. If you prefer just running the python files:
    a. Change directory in your terminal to the root of the “Coding_HW1_2020_Nix” folder.
    b. Run “python DL1_Splitting.py” to split the code.
    c. Run “python DL2_Training.py” to run the training. The results will be seen in the terminal.

Report

For the following assignment, the images given were split class-wise.

Doing the split in such a manner would make our training, validation and test sets more representative of real world data as each of these sets would have a more proportionate number of instances of each class. If we had randomly split the whole data instead, this could potentially cause the data to be skewed and imbalanced should, for example, a majority of a certain class had been assigned to the train set.

Using the class-wise average accuracy performance measure with each of the 4 Binary SVMs that were trained in a One vs All manner and then tested against the validation set, the best regularization constant c overall was c=0.01.

```
The class-wise average accuracy over the 4 seasons for c=0.01 is 0.3748
391018306636
The class-wise average accuracy over the 4 seasons for c=0.1 is 0.33964
83736515201
The class-wise average accuracy over the 4 seasons for c=0.316227766016
83794 is 0.34182177590715923
The class-wise average accuracy over the 4 seasons for c=1 is 0.3437014
751552795
```

```
The class-wise average accuracy over the 4 seasons for c=3.162277660168
3795 is 0.3489319426283099
The class-wise average accuracy over the 4 seasons for c=10 is 0.341821
77590715923
The class-wise average accuracy over the 4 seasons for c=10.0 is 0.3418
2177590715923
Overall, the best class-wise average accuracy was with c=0.01
```

Using this value of c=0.01, proceeded to train the 4 Binary SVMs again, this time with train and validation sets combined.

Results:

Class-wise averaged accuracy: `0.3993223993223993`
Vanilla accuracy: `0.7470588235294118`