***Machine Learning-*** *Homework 3*

Deadline: 2021/06/10

1. Consider the given data set with 118 inputs and one output. (a) **Binary Classification**: According to the guide line of HW1, please select a “nonlinear” machine learning model, e.g., neural network, SVM, random forest, …, to solve the binary classification of dataset. Details of your results (includes inputs, output, model, structure, and learning algorithm parameters, cross validation) should be given. (b) **Discussion of adding noise**: If noise with different values of SNR is added to the input and output respectively, please discuss its impact on classification results. (c) As above, add the noise to input and output respectively to observe the corresponding results in accuracy, convergence, etc. (d) **Classification:** Repeat part (a) in four categories classification, and compare the results of two methods to obtain the analysis.

**Note that the model (model, structure, activation function, …) and learning algorithm (loss function, learning algorithm, learning rate…) descriptions should be given and introduced.**

* Description of the data and the format of submission:

The folder has four .csv file containing training set and test set in binary classification and four categories classification. Each data contains 1 number (feature\_0), 118 features (feature\_1 to feature\_118) and 1 output (label). Table 1 and Table 2 presents the content of the data.

Please upload the report in pdf, and part (a), (c), and (d) need to upload code and model.

|  |  |
| --- | --- |
| Description of File | |
| File | Application |
| 2class\_trianing.csv | Binary classification for (a) and (c) |
| 2class\_test.csv |
| 4class\_trianing.csv | Four categories classification for (d) |
| 4class\_test.csv |

Table 1

|  |  |
| --- | --- |
| Table 2  Description of Data | |
| Column Name | Content |
| feature\_0 | Number |
| feature\_1 ~ feature\_118 | Input (X) |
| label | Output (Y) |