National Taiwan University **Introduction to Machine Learning and Deep Learning**

Department of Civil Engineering Instructor: C.-S. CHEN

Homework 3

Essay and Programming, Due 21:00, Wednesday, November 17, 2021

Late submission within 24 hours: score*0.9;

Late submission before post of solution: score*0.8 (the solution will usually be posted within a week); no late submission after the post of solution)

Total 90% + 30%(Bonus)

1. **(30%)** Name your file using student id, homework number and problem number scheme. For example, `b08501047_hw3_p1.pdf`. Consider a data set with three data points in $\mathbb{R}^2$:

$$\mathrm{X} = \begin{bmatrix} 0 & 0 \\ 0 & -1 \\ -2 & 0 \end{bmatrix} \quad \boldsymbol{y} = \begin{bmatrix} -1 \\ -1 \\ +1 \end{bmatrix}$$

   Manually solve the support vector machine problem to obtain the optimal hyperplane $(b^*, \mathbf{w}^*)$ and its margin.

2. **(60%)** Name your file using student id, homework number and problem number scheme. For example, `b08501047_hw3_p2.pdf`. Consider a dataset given in the following:

x	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
y	1	1	1	-1	-1	-1	-1	1	1	1

   Similar to Bagging, we can pick different training examples to obtain a new training set in AdaBoost. Supposed we have three rounds of boosting and each round has the following training record respectively:

   Boosting Round 1

x	0.1	0.4	0.5	0.6	0.6	0.7	0.7	0.7	0.8	1
y	1	-1	-1	-1	-1	-1	-1	-1	1	1

   Boosting Round 2

x	0.1	0.1	0.2	0.2	0.2	0.2	0.3	0.3	0.3	0.3
y	1	1	1	1	1	1	1	1	1	1

   Boosting Round 3

x	0.2	0.2	0.4	0.4	0.4	0.4	0.5	0.6	0.6	0.7
y	1	1	-1	-1	-1	-1	-1	-1	-1	-1

   We can use decision stump to classify the problem. The best split point for each round is:

Round	Split Point	Left Class	Right Class
1	0.75	-1	1
2	0.05	1	1
3	0.3	1	-1

Use AdaBoost and follow what we have covered in the toy example to compute ε_i, α_i and the updated weights. Finally find the combined classifier H.

3. **(Bonus 30%)** Name your file using student id, homework number and problem number scheme. For example, `b08501047_hw3_p3.pdf`. Consider the following classification data and module imports:

```
from sklearn.datasets import make_blobs
from sklearn.metrics import zero_one_loss
from sklearn.model_selection import train_test_split
import numpy as np
import matplotlib.pyplot as plt
from sklearn.ensemble import GradientBoostingClassifier

X_train, y_train = make_blobs(n_samples=5000, n_features=10,
centers=3, random_state=10, cluster_std=5)
```

Using the gradient boosting algorithm `GradientBoostingClassifier` with B = 100 rounds, plot the training loss vs. boosting rounds for γ = 0.1, 0.3, 0.5, 0.7, 0.9. Report your conclusions regarding the relation between B and γ and other findings you have learned from this exercise. (**hint**: the attribute `train_score_` of `GradientBoostingClassifier` allows you to access training loss of each round of boosting.)

- **Submission Format:** Please compress these pdf files into `yourStudentId_hw3.zip`, then upload it to NTU COOL.