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, Thursday, October 27, 2022

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 120%

If have any questions in this homework, contact TA 江郁瑄 (r10521605), and 李俊昇 (r11521608)

1. (30%) Name your file `split_tree.pdf`. Consider the training examples shown in the Table below for a binary classification problem.

Instance	a_1	a_2	a_3	Target Class
1	T	T	1.0	+
2	T	T	6.0	+
3	T	F	5.0	−
4	F	F	4.0	+
5	F	T	7.0	−
6	F	T	3.0	−
7	F	F	8.0	−
8	T	F	7.0	+
9	F	T	5.0	−

(a) Compute the entropy of this collection of training examples with respect to the class before splitting.

(b) Compute the information gains when the attributes a_1 and a_2 are used for splitting from the root. What is the best split (between a_1 and a_2) according to the entropy index?

(c) Similar to (a) and (b), compute the Gini index of this collection of training examples with respect to the class before splitting. Compute the gains from the Gini index when the attributes a_1 and a_2 are used for splitting from the root. What is the best split (between a_1 and a_2) according to the Gini index?

2. (30%) Name your file `hyperplane.pdf`. Consider the data below and a hyperplane $(b, \mathbf{w})$ that separates the data:

$$\mathbf{X} = \begin{bmatrix} 0 & 0 \\ 2 & 2 \\ 2 & 0 \end{bmatrix} \quad \boldsymbol{y} = \begin{bmatrix} -1 \\ -1 \\ +1 \end{bmatrix} \quad \mathbf{w} = \begin{bmatrix} 1.2 \\ -3.2 \end{bmatrix} \quad b = -0.5$$

(a) Compute $\rho = \min_{n=1,\cdots,N} y_n(\mathbf{w}^T\mathbf{x}_n + b)$

(b) Compute the weights $\frac{1}{\rho}(b, \mathbf{w})$ and show they satisfy $\min_{n=1,\cdots,N} y_n(\mathbf{w}^T\mathbf{x}_n + b) = 1$.

(c) Write a Python script to plot the data points and the hyperplane and compute the margin. Verify the margin is equal to $\frac{1}{\|\mathbf{w}^*\|}$ where * is the canonical representation of the hyperplane.

3. (30%) Name your file `ErrRate.py`. Consider an ensemble of 15 binary classifiers, each of which has an error rate of $\varepsilon_{\text{base}}$. The ensemble classifier predicts the class label of a test example by taking a majority vote on the prediction made by the base classifiers. Assume that the base classifiers are independent of each other and the ensemble makes a wrong prediction only if more than half of the base classifiers predict incorrectly. Write a python program to compute $\varepsilon_{\text{ensemble}}$ and plot the relationship of $\varepsilon_{\text{base}}$ vs. $\varepsilon_{\text{ensemble}}$.

Write a function called "output" to return the string as shown in sample.out , and a function called "draw" to draw and save the plot. Please turn .ipynb into .py and submit .py file named "`ErrRate.py`"

4. (30%) Name your file `adaBoost_1d.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 a 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.

- **Submission Format:** Please compress `split_tree.pdf`, `hyperplane.pdf`, `ErrRate.py` and `adaBoost_1d.pdf` into `yourStudentId_hw3.zip`, then upload it to NTU COOL.