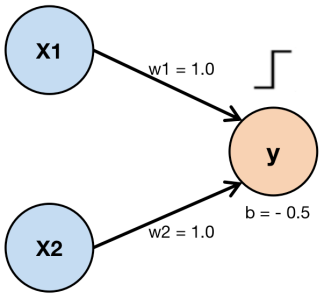
Assignment 3

Deadline: 23:55, December 11, 2023

**1.**Below is an example network for the **OR** operation. Given data points (0,0), (0,1), (1,0), (1,1), the network correctly generates labels 0,0,0,1.



|  |  |  |  |
| --- | --- | --- | --- |
| C:\Users\ADMINI~1\AppData\Local\Temp\ksohtml13132\wps4.jpg | C:\Users\ADMINI~1\AppData\Local\Temp\ksohtml13132\wps5.jpg |  | y (label) |
| 0 | 0 | -0.5 | 0 |
| 0 | 1 | 0.5 | 1 |
| 1 | 0 | 0.5 | 1 |
| 1 | 1 | 1.5 | 1 |

Please construct a simple Artificial Neural Network to implement three-input operations for AND Operation. The corresponding input-output pairs are given in the table below.

|  |  |  |  |
| --- | --- | --- | --- |
| C:\Users\ADMINI~1\AppData\Local\Temp\ksohtml13132\wps1.jpg | C:\Users\ADMINI~1\AppData\Local\Temp\ksohtml13132\wps2.jpg | C:\Users\ADMINI~1\AppData\Local\Temp\ksohtml13132\wps3.jpg | y (label) |
| 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 0 |
| 0 | 1 | 0 | 0 |
| 0 | 1 | 1 | 0 |
| 1 | 0 | 0 | 0 |
| 1 | 0 | 1 | 0 |
| 1 | 1 | 0 | 0 |
| 1 | 1 | 1 | 1 |

You have to **describe its network structure**, and **how does your ANN predict the True labels**. (Same way the OR example presented) In addition, please elaborate what loss function you would choose for the training process. You do not need to actually train the parameters, and fewer parameters will be preferred. (25 marks)

**2.**For a binary classification problem, suppose we use an ensemble model with three base learners. The three base learners are independent of each other and each base learner has an error rate ε (errors are uncorrelated). Then compute the error rate of the ensemble model in terms of ε. When ε=0.3, compute the value of the error rate of the ensemble model. (25 marks)

**3.** Please train a model using a medical dataset for diabetic prediction.

You will be provided with 2 csv files training dataset (train) and test dataset (test), as the diabetes column being the label column. (50 marks).

Each line in both Diabetes\_trainingset and Diabetes\_testset is a data sample, and all features are separated with a comma.

Based on the training set, you need to train a model out of the following candidate models: decision tree, nearest neighbor classifier, support vector machine, neural network as well as ensemble models.

You need to predict the class **label (Credit\_Score)** of the test data points in the Diabetes\_testset. The prediction of your model should be stored in a csv file. **Each line** of the file stores the predicted class label for the corresponding test data point.

You need to submit a **report** to describe the models you used, operations you made, and the **corresponding code**.

Grading will be based on the testing accuracy (70%) and the writing of the report (30%).

You need to submit the source code, your predictions (in a csv file), and the report (in a pdf file) before the deadline via Blackboard.