Jacobs University Bremen
SHOROUK G. AWWAD

11,05, 2019

TERAKI Quiz 1

Problem 1.1

Solution:

Given that the weight vector $(W_0, W_1) = (30, 30)$, the bias $b = -20$, and the input is binary $X_in = (0, 1)$ and the output should be of the same binary form, the $Sigmoid$ function has been decided on to be the activation function.

So,

$f(x) = (W_0X_0 + W_1X_1 + b)$ and $g(f) = 1/(1 + e^{-20})$

By drawing a truth table to all possible inputs for X_0 and X_1:

X_0	X_1	g(f)	g̃(f)
0	1	0.9998	1
1	0	0.9998	1
0	0	2.06e-9	0
1	1	0.99999	1

from the above table, we notice that these Neural networks are approximating **OR** function.

Problem 1.2

Solution:

a) for balanced data, things will be so simple , so Metric models like **Accuracy** will work fine and will give accurate results.

b) for Imbalanced datasets, some Metrics become misleading more than helping, since one class can have the majorty of the data distribution compared to some other classes, which makes some shallow algorithms throw most of the training data in this class, except for a litte number with huge deviation.
So in this case, if metrics like **Accuracy** have been used - like to determine the number of right predictions in total realtive to the number of predictios, without considerng the classes weights, this will lead to high accuracy but wrong berformance if the model tries to predict differently structured data.
So, the best choice in this case would be :

* True Positive Rate
* True Negative Rate
* False positive Rate
* Percision
* F1 score
* Mcc

Also, one way to solve this issue is by controling the number of samples. for example : increasing the number of samples in the minor classes or cutting off a random number of samples from the major class.

Problem 1.3

Solution:

a) The used average sample rate is 51.26 Hz.
Please refer to the sample Rate.ipynb

for b and C, please refer to the code files and the documentation attached