Exercise 10 of Machine Learning [IN 2064]

Name: Yiman Li
Matr-Nr: 03724352
cooperate with Kejia Chen(03729686)

Problem 1

Failed to accomplish the total mathematical task! Can't tell the difference between condition a and b!

One of the nice properties of logistic regression is that the logistic cost function (or max-entropy) is convex, and thus we are guaranteed to find the global cost minimum. But, once we stack logistic activation functions in a multi-layer neural network, we'll lose this convexity, and so in practice local maxima of the likelihood may be found, corresponding to local minima of the error function, see in [1]. So normally, we shall say that the linear regression loss $\mathcal{L}_{LS}(\boldsymbol{w}_{LS}^*)$ is smaller than the neural nerwork loss $\mathcal{L}_{NN}(\boldsymbol{W}_{LS}^*)$.

Problem 2

Failed to accomplish the total programming task!

The unfinished version to be seen in the end.

References

[1] Christopher M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer, 1 edition, 2007.