ML in Fundamental Physics

Ludwig-Maximilians-Universität München – Dr. Sven Krippendorf

---

Exercise Sheet 3

Submission by Thursday 14$^{\text{th}}$ 17:00 via Moodle.
Corrections due by Monday 18th$^{\text{th}}$ 16:00 via Moodle.
Format for all questions: individual Jupyter notebooks .ipynb.
Text for question 3 in markdown in the Jupyter notebook.

1 MNIST – regression (content 3 points)

The aim of this exercise is to get some experience with logistic regression and regularisation. We shall use the implementations in sklearn for this exercise.[1]

- Using the SGCClassifier and the LogisticRegression, build classifiers for the MNIST dataset.
- Find some handwritten digits which your classifiers are not capable of classifying correctly. How do they differ from correctly classified digits?
- Scan over hyper-parameters (i.e. l1, l2 regularisation, different optimisers) and compare the performance.
- Compare your results to results you would have obtained by pure random guessing.

2 1D Ising – Regression (content 2 points)

In the lecture we have shown different performances of linear regression on predicting the energy in the 1D Ising model depending on our choice of regularisation. The aim of this exercise is to analyse this performance a bit further.

- Instead of the built-in scoring function, plot the error E_{out} for the choice of parameters used in the lecture.
- How does the performance change when changing the number of samples?

3 Perceptron (XOR) and Shannon information content (content 2 points + 1 bonus point)

- Show that a single perceptron does not have the capability to realise the XOR function.
- $\star$ How can this be avoided using multiple perceptrons?
- For independent random variables ($P(X,Y) = P(X)P(Y)$) show that Shannon entropy is additive, i.e. $H(X,Y) = H(X) + H(Y)$.
- $\star$ This is slightly off-topic but interesting nevertheless. Utilising information content (e.g. by making it explicit on your decision tree), think about a solution for the following problem. You are given 12 balls, all equal in weight except for one that is either heavier or lighter. You are given a two-pan balance to use. In each use of the balance you may put any number of 12 balls on the left pan, and the same number on the right pan. Your task is to design a strategy to determine which is the odd ball and whether it is heavier or lighter than the others in as few uses of the balance as possible.

[1] Documentation for sklearn can be found online and it is useful to get familiar with the way this is written.