Short report on lab assignment 1

Learning and generalisation in feed-forward networks — from perceptron learning to backprop

Ivan Stresec, Frano Rajic

September 11, 2020

Please be aware of the constraints for this document. The main intention here is that you learn how to select and organise the most relevant information into a concise and coherent report. The upper limit for the number of pages is 6 with fonts and margins comparable to those in this template and no appendices are allowed.

These short reports should be submitted to Canvas by the authors as a team before the lab presentation is made. To claim bonus points the authors should uploaded their short report a day before the bonus point deadline. The report can serve as a support for your lab presentation, though you may put emphasis on different aspects in your oral demonstration in the lab. Below you find some extra instructions in italics. Please remove them and use normal font for your text.

1   Main objectives and scope of the assignment

List here a concise list of your major intended goals, what you planned to do and what you wanted to learn/what problems you were set to address or investigate, e.g.
Our major goals in the assignment were

- to investigate the performance of classification single-layer perceptrons (SLPs) in various settings

- to observe the capabilities of MLPs in the context of solving non-linearly separable patterns in both classification and regression problems

- to implement MLPs in a chaotic time-series prediction[1]

---

[1] FIX THIS ONE

Then you can write two or three sentences about the scope, limitations and assumptions made for the lab assignment

2 Methods

Mention here in just a couple of sentences what tools you have used, e.g. programming/scripting environment, toolboxes. If you use some unconventional method or introduce a clearly different performance measure, you can briefly mention or define it here.

For the first assignment we used Python as a programming language and its library Numpy for data representation and matrix operations. For graphs we used the Matplotlib library. In the second assignment we used Pytorch, as well as the aforementioned libraries. As an IDE, we've used JetBrains' PyCharm for both assignments.

3 Results and discussion - Part I

*Make effort to be **concise and to the point** in your story of what you have done, what you have observed and demonstrated, and in your responses to specific questions in the assignment. You should skip less important details and explanations. In addition, you are requested to add a **discussion** about your interpretations/predictions or other thoughts concerned with specific tasks in the assignment. This can boil down to just a few bullet points or a couple of sentences for each section of your results. Overall, structure each Results section as you like, e.g. in points. Analogously, feel free to group and combine answers to the questions, even between different experiments, e.g. with linearly separable and non-separable data, if it makes your story easier to convey.*

*Plan your sections and consider making combined figures with subplots rather than a set of separate figures. **Figures** have to condense information, e.g. there is no point showing a separate plot for generated data and then for a decision boundary, this information can be contained in a single plot. Always carefully describe the axes, legends and add meaningful captions. Keep in mind that figures serve as a support for your description of the key findings (it is like storytelling but in technical format and academic style.*

*Similarly, use **tables** to group relevant results for easier communication but focus on key aspects, do not overdo it. All figures and tables attached in your report must be accompanied by captions and referred to in*

3.1   Classification with a single-layer perceptron *(ca.1 page)*

Combine results and findings from perceptron simulations on both linearly separable and non-separable datasets. Answer the questions, quantify the outcomes, discuss your interpretations and summarise key findings as conclusions.

In solving classification problem using an SLP one can easily observe its efficiency and its limitations.

Firstly, we have implemented the perceptron rule (implying a sequential learning scheme) and the Delta learning rule (batch learning) to a randomly generated, linearly separable classification dataset. Both have performed well and converged quickly...

Secondly, we have tested the differences between batch and sequential learning for the Delta rule SLP. In terms of epochs, both approaches converge relatively quickly, with ... being somewhat quicker

Thirdly, we have tested removing the bias term with the Delta rule in batch mode. Even without testing it, it is clear that the algorithm can converge only when the data is linearly separable by a line which goes through the center of the coordinate system.

3.2   Classification and regression with a two-layer perceptron *(ca.2 pages)*

3.2.1   Classification of linearly non-separable data

It seems that one (decision boundary) or two plots (inclusing learning curves) should suffice. Build a story around the questions in the assignment. Include concise motivation for your findings and potential interpretations/speculations.

3.2.2 The encoder problem

Here you do not really need any illustrations, this could be a very short section reporting on your experiments in line with the assignment questions.

3.2.3 Function approximation

This subsection requires plots to reflect intuitive visual interpretation of the results. Make sure that you condense information and avoid any excessive plotting. Here you might also need to incorporate some illustration of the network's generalisation performance or use a table to systematically report the results requested in the assignment.

4 Results and discussion - Part II *(ca.2 pages)*

Here you do not have to introduce the problem or define Mackey-Glass time series, as you should focus on the results. You could divide them into two parts as the following two suggested subsections but you might as well keep your story under the main heading of Part II of the assignment. Importantly, always clearly state what network architecture you use, crucially with the number of hidden nodes, systematically report average results with various manipulations (regularisation etc.) and pay attention to differences between training, validation and test errors. Illustrating the outcome of your network predictions along with the original chaotic time series can also be very helpful. Finally, since you compare two- and three-layer architectures, make sure that you do not jump to any conclusions based on a small number of simulations unless you have statistically convincing evidence (when you comare the mean performance measures, their second moment is also relevant). In this part it may be particularly desirable to rely on tables.

4.1 Two-layer perceptron for time series prediction - model selection, regularisation and validation

4.2 Comparison of two- and three-layer perceptron for noisy time series prediction

5 Final remarks *(max 0.5 page)*

Please share your final reflections on the lab, its content and your own learning. Which parts of the lab assignment did you find confusing or not necessarily

helping in understanding important concepts and which parts you have found interesting and relevant to your learning experience?
Here you can also formulate your opinion, interpretation or speculation about some of the simulation outcomes. Please add any follow-up questions that you might have regarding the lab tasks and the results you have produced.