UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

COS711 Assignment 2

Supervised Neural Network Training

Due date: 20 April 2020, at 23h30

1 General instructions

You have to submit a pdf document, containing a technical report wherein you describe what you have done, present and discuss your findings. Guidelines for writing your report are provided in this specification document. The report will be checked for plagiarism using the Turnitin system, and should be submitted through the ClickUp system. You are advised but not required to typeset your report in LaTeX.

2 Supervised Learning

For this assignment, you will compare passive learning to curriculum learning of feedforward neural networks (NNs), and implement an architecture selection algorithm of choice. NNs will be used to predict the quality of wine.

2.1 Data set

The dataset for this assignment can be obtained from the UCI Machine Learning Repository: `https://archive.ics.uci.edu/ml/datasets/Wine+Quality`. You can also download the data from the CS website. The dataset contains 11 input variables, and 1 output variable in the integer range `{0,...,10}`. Two datasets were created by the authors, using red and white wine samples. The inputs include objective tests (e.g. PH values) and the output is based on sensory data (median of at least 3 evaluations made by wine experts). Each expert graded the wine quality between 0 (very bad) and 10 (very excellent). Your task is to construct a NN to make a prediction about the quality of wine based on the input characteristics.

2.2 Your task

Your task is to train a NN to predict the wine quality. Note that the task can be interpreted either as a regression problem (output a real number between `0` and `10`), or as a classification problem (10 discrete classes corresponding to the qualities of wines). It is up to you to decide whether you want to do classification or regression. Whatever choice you make, try to justify it with a relevant discussion in your report.

For the purpose of this assignment, you will need to perform the following steps: pre-process the data; optimise NN parameters; implement an architecture selection algorithm (besides regularisation); compare passive learning to curriculum learning.

2.2.1 Data preparation

The given data contains numeric attributes that lie in various ranges. Analyse the data set, and pre-process it in a way that will make it possible for the NN to effectively discover the hidden relationships between inputs and outputs. The authors suggest that some inputs may be irrelevant. Try to verify this in your data analysis.

Note that, if you choose to do the classification task, then you will be dealing with class imbalance. Imbalance is common is real-world datasets. You must decide how to deal with this imbalance, and discuss the chosen approach in the report. For the regression task, discuss your approach to outliers.

2.2.2 Neural Network Parameter Optimisation

As discussed in class, the performance of your NN model greatly depends on various hyperparameters, such as the number of layers and hidden units, activation functions, error function, optimisation algorithm parameters, etc. You will have to choose the hyperparameter values for your NN model. Your report **must** contain a section justifying all hyperparameter choices. Two justifications are acceptable: (1) theoretical insight; (2) empirical evidence. I.e., if you cannot decide on a value for a certain hyperparameter analytically, you have to run some experiments to see which value performs better than others.

You must empirically compare at least two different hyperparameters of your choice.

2.2.3 Implementing Architecture Selection

Architecture selection is a difficult task, and numerous approaches were suggested in literature to automate it. For this assignment, you must pick either a constructive (i.e. ability to add neurons/layers) or a pruning (i.e. removal of unnecessary parameters) approach. Note that regularisation does not count as architecture selection, since it does not actually remove unnecessary parameters, but simply drives them to zero. Refer to the lecture slides for a discussion of various architecture selection approaches. You are also welcome to find alternative sources, or come up with your own approach.

Perform a comparison between a NN with a static architecture, and a NN with an architecture selection strategy. Which one performed better/converged faster? Discuss all your results thoroughly. Remember that simply pasting a table with numbers will not yield any marks. If you see that one approach is doing better than the other, give a hypothesis for why it is the case.

2.2.4 Passive VS Curriculum Learning

Finally, perform a comparison between passive learning and curriculum learning. Note that curriculum learning evaluates data patterns based on their difficulty, and starts with easier patterns before harder patterns are shown to the NN. You can pick any existing curriculum learning approach, or propose your own.

Discuss your results thoroughly. Remember that simply pasting a table with numbers will not yield any marks. If you see that one approach is doing better than the other, give a hypothesis for why it is the case.

3 Notes

- Implementation
  - You may use any programming language and platform
  - You may use a neural network library/framework
  - Two datasets are given: for red and white wine. It is up to you how the two are used. You can merge them into a single dataset, or reserve the smaller of the two for generalisation purposes. The second approach will reveal whether the characteristics of one wine apply (generalise) to another wine.
- Report
  - You must report on all data preparation steps taken
  - You must report on all algorithm parameters used, and substantiate your choices
  - Training, generalisation, and classification (if applicable) errors have to be reported. When classification error is reported, clearly indicate what part of the dataset was used to generate it. Remember to report means with the corresponding standard deviations.

4 Marking and general guidelines

For this assignment you have to submit a research report where you discuss your findings. Your reports must follow the IEEE conference format (`http://www.ieee.org/conferences_events/conferences/publishing/templates.html`). You may use the Latex or the Word template, however it will serve as good academic writing practice to utilise LaTeX. There is also a strict page limit of **8 pages** for this assignment. Given the imposed two column format it would require a substantial amount of writing to exceed this limit.

This is not a course in technical and report writing; however, you should at least attempt to follow some accepted document writing techniques and make your report as readable as possible. You are more likely to obtain a higher mark if your report generates a good impression with the marker and is void of general errors like spelling and grammar mistakes.

A typical report would consist of the following sections:

1. **Abstract**

   The abstract should briefly summarise the purpose and findings of the report.

2. **Introduction**

   The introduction sets the stage for the remainder of your report. You usually have very general statements here. The introduction prepares the reader for what to expect from reading your report. In general, the introduction should either contain or be a summary of your ENTIRE report. Keep the introduction concise, try to limit it to 1 page maximum.

3. **Background**

   A very high level discussion on the problem domain and the algorithms and/or approaches that you have used. Do not be too specific on the algorithms and approaches. This section is typically where the “base cases” of concepts that appear throughout the remainder of your report are discussed. It is also an ideal place to refer a reader to other sources containing relevant information on the topic which is outside the scope of your assignment. Remember to discuss very generally. After reading this section the marker should be able

to determine whether or not you know what you're talking about. Try to limit this section to 1 page maximum.

4. **Experimental Set-Up**

   In this section you discuss how you approached, implemented and solved your assignment. Mention the values set for the algorithm's control parameters, how many simulations you have run and what the characteristics for candidate solutions to your problems are. After reading this section (in addition to the background) the reader should be able to duplicate your experiments to obtain similar results to those obtained by you. This is also the section where your discussion specialises on the concepts mentioned in the background section. Be very specific in your discussions in this section.

5. **Research Results**

   This is the section where you report your results obtained from running the experiments as discussed in the experimental set-up section. You have to give, at least, the averages and the standard deviations for all the experiments/simulations. Training, generalisation, and classification errors (if applicable) have to be reported. Thoroughly discuss the results that you've obtained and reason about why you obtained the results that you have. Answer questions like "are these results to be expected?" and "why these results occurred?" and "would different circumstances lead to different results?"

6. **Conclusion(s)**

   Very general conclusions about the assignment that you have done. This section "answers" the questions and issues that you've raised and investigated. This section is, in general, a summary of what you have done, what the results were and finally what you concluded from these results. This is the final section in your document so be sure that all the issues raised up until now are answered here. This is also the perfect section to discuss what you have learnt in doing this assignment.

Please **do not** include any code or pseudocode in the report, unless you are proposing a novel algorithm. Research reports must focus on the scientific contributions. We just assume that you can code – you do not have to prove it anymore!

4.1 Marking

The following general breakdown will be used during the assessment of this assignment:

Category	**Mark Allocation**
Report Structure	5 marks
Background	5 marks
Experimental Setup	10 marks
Data Preparation	10 marks
Parameter Optimisation	10 marks
Architecture Selection	25 marks
Curriculum Learning	25 marks
Conclusions	5 marks
References	5 marks
TOTAL	100 marks

Submit only the PDF report. No additional files of any sort should be submitted. Upload the PDF file to the appropriate assignment upload on ClickUp. Multiple uploads are allowed, but only the last one will be marked. The deadline is **20 April 2020, at 23h30** .