UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

COS711 Assignment 2
Supervised Neural Network Training

Due date: 25 October 2021, 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 the performance of gradient descent with that of conjugate gradient descent, and implement a curriculum learning approach. NNs will be used to predict astronomical objects based on recorded physical observations.

2.1 Data set

The dataset for this assignment can be obtained from the Kaggle website: `https://www.kaggle.com/lucidlenn/sloan-digital-sky-survey`. You can also download the data from the CS website. The dataset contains 10,000 observations, each observation comprising 17 input variables, and 1 output variable in the discrete range: {STAR, GALAXY, QUASAR}. Your task is to construct a NN to make a prediction about the type of astronomical object observed based on the given input values.

2.2 Your task

For the purpose of this assignment, you will need to perform the following steps: pre-process the data; optimise NN parameters; implement gradient and conjugate gradient descent learning; 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.

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 Conjugate Gradient Descent

Conjugate direction methods can be regarded as lying between the method of the steepest descent (first-order method that uses gradient) and Newton's method (second-order method that uses the Hessian matrix). As discussed in class, line search is used to determine the step size along the direction conjugate with the direction of the gradient.

For this task of the assignment, you can choose any variation of the gradient and conjugate gradient descent methods. Perform a comparison between a NN trained with a gradient descent variant, and a NN trained with a conjugate gradient descent variant. 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. Remember that comparing a single run of each algorithm is invalid: perform multiple independent runs, and compare means together with the corresponding standard deviations.

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. Remember that comparing a single run of each algorithm is invalid: perform multiple independent runs, and compare means together with the corresponding standard deviations.

3 Notes

- Implementation
    - You may use any programming language and platform
    - You may use a neural network library/framework
- 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 (on the test set) 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
Conjugate VS Gradient Descent	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.