How to Write a Thesis in the ML group

Tobias Uelwer, Stefan Harmeling

September 3, 2020

1 Find a Topic

Finding a topic is a time-consuming and iterative process. If you have your own idea for a research project, great! In that case, let us know what you have in mind.

If you do not have an idea already, do not worry! A good starting point are the large machine learning conferences where the newest reseach papers get published:

1. Conference on Neural Information Processing (`https://neurips.cc`)
2. International Conference on Machine Learning (`https://icml.cc`)
3. International Conference on Learning Representations (`https://iclr.cc`)
4. International Conference on Artificial Intelligence and Statistics (`https://www.aistats.org`)

All these conferences publish their proceedings, i.e., the collection of all published papers in that year, online without a pay-wall. For your thesis you can consider the last ten years, so there is a huge load of papers you could select from.

If you found an interesting paper, try answering the following questions.

1. What is this paper about and to which sub-field does it belong?
2. What are necessary backgrounds that you need to explain in your thesis so that the reader can understand your thesis without doing a whole literature review by herself/himself?
3. How can you extend the experiments or the model in your thesis? This is probably the most important point, so try to find a twist which makes your thesis more interesting! You could for example change some parts of the model in a way that makes sense or you could do some experiments that were not done in the paper, e.g., by trying to use a different dataset or by analysing an aspect that was not considered in the original work.

After that you could make an appointment to discuss your project with us.

It is also possible to write a survey on a broader topic covering multiple different papers. But if you decide to write a survey keep in mind that the writing of your thesis will be examined more critically as you usually do not include your own results in your thesis.

2 Write an Abstract

The next step is writing an abstract which briefly, i.e., roughly one page, summarizes your plans for the project. Start with introducing the sub-field of the paper you are considering in 2-3 sentences. After that explain the idea of the original paper in 3-4 sentences. At the end explain how you plan to extend the model/experiments in 5-6 sentences.

Important: The abstract is not a contract! It is also fine to change the scope of your thesis while you are writing! Just let us know, e.g., by writing a brief email.

3 Registering Your Thesis

Send the abstract to Prof. Harmeling and to Tobias Uelwer:

`harmeling@hhu.de, tobias.uelwer@hhu.de`

If Prof. Harmeling gives his ok, you can register your thesis at:

`https://studierende.hhu.de`

The registration usually takes a couple of days. After that your time is running! For a Bachelor's thesis your deadline is 3 months after the registration. For a Master's thesis it is 6 months after registering the thesis.

4 Give a Talk About Your Thesis

A couple of weeks after registering your thesis we would kindly ask you to give a short talk, i.e., 30 minutes, about the topic of your thesis within our Oberseminar. The Oberseminar is meant to be a forum for the all interested students where questions can be asked and problems can be discussed. In your talk you can introduce the other students to your topic, explain what you are planning to explore and give an overview of the problems you are currently working on. If there is some concept in the original paper you did not understand yet you can also address this within your talk and ask the other participants how they understood it! You do not have to present your final results in this talk. If you have preliminar results you can of course present them in your talk.

The Oberseminar is organized using Mattermost which is the chat platform of the computer science departement. You can access it through

`https://mattermost.cs.uni-duesseldorf.de`

after registering at:

`https://tofu.cs.uni-duesseldorf.de`

Use your university login to access the page. Make sure you check the box for GitLab/Mattermost. On the Mattermost server join the ML team and the Oberseminar channel. (The teams can be selected on the very left. Channels can be joined by clicking on the small '+' next to the public channel list.)

5 Write the Thesis

A good resource on mathematical writing for computer scientists is Knuth et al. [1] available at `https://jmlr.csail.mit.edu/reviewing-papers/knuth_mathematical_writing.pdf`.

Your thesis should be type-setted using LaTeX. You can, e.g., use the thesis template provided by Prof. Conrad at `https://dbs.cs.uni-duesseldorf.de/lehre/bmarbeit/vorlage_2020_01_09.zip`.

We would also like to highlight some very common mistakes below.

5.1 Common Mistakes

5.1.1 Capitalization

In English each word of the title of a thesis, a section or a sub-section - except for small words like 'the' or 'a' - is usually capitalized. This is referred to as title case. Everything else is very wrong! If you are unsure about the capitalization of your title use the following online tool: `https://titlecaseconverter.com`.

5.1.2 Equation Numbering

Always put equation numbers, since even you might not refer to a particular equation, your reviewer might do it.

6 Hand in Your Thesis

6.1 Your Thesis

You should hand your thesis in by uploading a PDF version of it (produced by LaTeX) to `https://studierende.hhu.de`. This will be the offically graded version of your thesis, which is also to be uploaded before your deadline. Furthermore please also hand in a softcover print version of your thesis, either by

giving it to Claudia Forstinger or Prof. Harmeling. This should happen two weeks after your deadline at the latest.

You do not have to sign the Erklärung at the beginning of your thesis. It is sufficient that you check the corresponding box when handing in the thesis online. However, most students include a signature (either scanned or via a tablet or paint) in the pdf version of the thesis. But this is optional!

6.2 Supplementary Material and Code

Include relevant supplementary material for your thesis (i.e. code implementing models or additional plots and tables) as an appendix to to your thesis. If you have a lot of code, consider printing it in a small but readable font size to save ink and paper. You should also consider uploading it to a GitHub repository, or a university internal GitLab repository (`https://git.hhu.de`) so we can run the code on our computer.

7 Get a Grade

We want to be transparent about the grading of your thesis. Here are some rule-of-thumbs:

1. If work out a simple paper you can get a grade up to 2.0.
2. Work out a tough paper will give you a grade up to 1.3.
3. If you find and discuss a creative extensions within your thesis your grade can be better up to 1.0.
4. If your thesis has a lot of spelling and comma mistakes, your grade will be up to 1.0 worse.

8 Ask Us

After having read up to this point, of course, you are welcome to ask us any open questions. We look forward to work with you!

Furthermore, if you spot any mistakes within this document or you have the feeling that there is something important missing, please send me an email (`tobias.uelwer@hhu.de`).

References

[1] Donald Ervin Knuth, Tracy Larrabee, Paul M Roberts, and Paul M Roberts. *Mathematical writing*, volume 14. Mathematical Association of America Washington, DC, 1989.