Master thesis – introduction – Fall 2024

Frank Miller, using slightly adapted slides from Oleg Sysoev

Statistics and Machine Learning
Deparment of Computer and Information Science
Linköping University

Admission requirements – the hard facts

- You must have passed at least 60 ECTS credits in programme courses, incl. machine learning (732A99) completed and 6 ECTS from semester 3

- All compulsory courses needed for writing the thesis, but the threshold requirements are relaxed
  - Tell me (Frank) when you have questions about that

Master thesis

- **… is a course → has start and end dates, mandatory sessions and deadlines**

Date	Activity
August 9	Deadline for sending in project descriptions to Frank.
August 14	Approval of thesis topic.
August 20 or earlier	Appointment of supervisor.
August 21	Introductory meeting at 10.15 in John von Neumann, B building.
September 9	Submitting the thesis proposal presentation to LISAM.
September 13	**Thesis proposal seminar. In John von Neumann, B building.**
October 18	Submitting the midterm presentation and supporting report to LISAM.
October 24	**Mid-term seminar. In John von Neumann, B building.**
November 22	Sending a thesis draft to the supervisor unless the supervisor states otherwise.
November 29	Decision is made by the supervisor regarding whether the thesis has good chances to be passed within the course time frames or the thesis work needs to be stopped and grade F given.
December 4	Sending a thesis draft to the opponent, examiner and supervisor.
December 11, 12 or 13	**Revision meeting. Mandatory, but date and place are decided in agreement between the student, opponent, supervisor and examiner.**
January 5, 2025	Submission of the close-to-final thesis to LISAM. *This version will be controlled for the plagiarism.* Send it also to the supervisor, examiner and opponent by email.
January 12, 2025	Submission of the presentation for the oral defense to LISAM.
January 13, 2025	**Oral defense seminar. In John von Neumann, B building.**
January 20, 2025	Last day to submit final version of the thesis for it to be reported within this semester.
January 27, 2025	Examiner's meeting where the grades are decided and then reported.
March 21, 2025	Last attempt for submitting the final version of the thesis (for students who become delayed; the maximum possible grade is C).
June, 2025	Second opportunity for the oral defense seminar (for students who become delayed).

- Mandatory to attend **all sessions marked in bold** (thesis proposal seminar, mid-term seminar, revision meeting, defense seminar)

Master thesis

- Thesis project is an individual work
- Missing some mandatory session may lead to postponing the examination of your work to next term! This may imply a new supervisor and possibly a new project

- **Note: total thesis writing time is very limited; all experiments done + thesis written in slightly more than 3 months**
  - 1-2 methods in depth rather than scratching many on the surface
  - Grade is not related to the accuracy of your model

- **Your thesis is formally in Statistics** → Important to make sure that Statistics is clearly there, e.g.,
  - Statistical methods as a comparison benchmark
  - Statistical evaluation measures
  - Hypothesis testing, prediction intervals, confidence intervals

Supervision

- The normal frequency of meetings with the supervisor is every second week
- At each meeting it is recommended to agree on what the next steps are

- It is mainly the student who sets "the agenda" for the meeting, i.e. you are expected to prepare questions and regularly **demonstrate the progress**
  - Start writing very early!
  - Submit manuscript parts to the supervisor in due time before a meeting (to facilitate feedback)

- Supervisor is doing examination as well, i.e. evaluates your
  - Progression
    - the supervisor has a right to request the texts demonstrating your progress at any time point!
  - Independence
    - Independence does not mean that you need to plan everything yourself and then come to the supervisor with the final results of your experiments!
    - You should be the driving force: show initiative, give suggestions regarding ways to solve the research problems
  - Communication and collaboration skills
- Supervisors can be experts in ML but not in statistics, or in statistics but not ML → take more own initiative around questions outside your supervisor's expertise; discuss potential solutions with your supervisor

Part 1: Problem formulation and literature review

August 21	**Introductory meeting at 10.15 in John von Neumann, B building.**
September 9	Submitting the thesis proposal presentation to LISAM.
September 13	**Thesis proposal seminar. In John von Neumann, B building.**

- **Time frame: Now – thesis proposal seminar**

- Make a careful **review of the background** to your problem, **available data** and the **specific questions** put by the company/internal supervisor
- Put your work in a scientific context
  - Has anyone studied your kind of problem previously? Which methods were used?
  - Are there studies of similar problems?
  - What lessons can be learnt?
  - What is new in your study?
  - Why did you choose a certain model/method/algorithm? Which are alternatives?

- Put focus on the literature review → Google, Google scholar, Google…
- Suggest potential methods to solve your research questions and discuss them with the supervisor
- Prepare for your thesis proposal seminar (oral with slides)

Thesis proposal seminar

- Each student presents his/her thesis topic
  - Data description
  - Background information and research questions
  - Potential methods to solve the research questions
- 15 mins presentation + 10 mins questions
- You are not supposed to have a completely defined strategy/plan for your thesis work at this stage
- Why putting on LISAM in advance?
  - Will be checked by the examiner → Better feedback opportunities

Thesis proposal seminar

- Thesis proposal should help you with your research questions → completely OK if you don't know how to optimally solve all stated research questions → give space for improvement suggestions from us in your talk as appropriate:
  - ”I propose to use this method, but I am not sure whether this one is the optimal …”
  - ”These approaches might be suitable but I'm not sure which one is better to use …”
  - ”Me and my supervisor are still not sure about how to approach this subproblem …”

- Clear explanation + informative presentation = more feedback

- Expected outcomes:
  - Each course attendant, supervisor and examiner should have been informed about and understood the objectives of the thesis projects and possible ways to achieve these objectives
  - A feedback on the research planning is obtained by the course attendants when appropriate

- **Mandatory to attend all sessions!**

Part 2: Main work, first phase

October 18	**Submitting the midterm presentation and supporting report to LISAM.**
October 24	**Mid-term seminar. In John von Neumann, B building.**

- **Time frame: Thesis proposal seminar – mid-term seminar**
- In the beginning:
  - Take the feedback from the proposal seminar into account
  - Sketch a time plan for your research (work packages)
  - Write introduction, literature review and objectives
- Work with your project (simulation, reflection, discussion, documenting)
- Organize a regular supervision, discuss intermediate results/problems with the company/internal supervisor

Part 2: Main work, first phase

- Write draft texts containing results and their analysis
  → **send/show your texts to the supervisor regularly**

- By the end of the phase, **at least 60%** of the work needs to be done, including the thesis writing part

- Ends by the mid-term seminar

Mid-term seminar

- Each student presents the progress so far. 20 minutes per thesis + 10 mins questions.
  - Train yourself to fit these time frames in advance
- A two-page written summary of the progress + sample thesis text, for example a completed piece of introduction section (in the appendix) + presentation slides should be uploaded to LISAM in time
- Expected outcomes: Each course attendant and the supervisor/examiner should become informed about what has been done so far and what is left to do for each thesis project
- Don’t assume that people remember what you explained in the thesis proposal
- Student’s benefits: You may get a critical feedback on your intermediate results and improvement suggestions
  - Make presentation clear and easy to understand, fit in time!
- **Mandatory to attend all sessions!**

Part 3: Main work, second phase

Time frame: Mid-term seminar – revision meeting

November 29	**Decision is made by the supervisor regarding whether the thesis has good chances to be passed within the course time frames or the thesis work needs to be stopped and grade F given.**
December 4	Sending a thesis draft to the opponent, examiner and supervisor.
December 11, 12 or 13	**Revision meeting. Mandatory, but date and place are decided in agreement between the student, opponent, supervisor and examiner.**

- Take aspects from the mid-term report seminar into account
- Work with your project (simulation, reflection, discussion, documenting)
- Parts of the thesis text need to be sent to the supervisor regularly
- Write draft pages containing results and their analysis
- Organize a regular supervision

Part 3: Main work, second phase

- Prepare the skeleton of your final thesis and include the parts that are already written → obtain a thesis draft
- Some parts (e.g., conclusions, discussions) can be missing but the draft needs to be complete enough for a supervisor to make decision
- Send this draft to the supervisor unless you got different instructions from the supervisor
- **Decision is made by the supervisor** regarding whether the thesis has
  - good chances to be passed within the course time frames, or
  - the thesis work needs to be stopped and grade F given.
- If failed, you may start new thesis work based on new project in the next term

Part 3: Main work, second phase

- Work on your thesis and complete the remaining parts of your draft:
  - Finalize your analyses and complete the results part of your thesis
  - Finalize the methodology chapter and the reference list of your thesis
  - Put strong effort on writing a good discussion part, consider limitations of your approaches, methods and algorithms – it should comprise several pages
  - Formulate your conclusions
    - One paragraph of conclusions = One research objective
  - **Write "Ethical considerations" section!**
  - **If you used generative AI supporting your thesis work, write a section about how and to what extend you have used it!**

- A thesis draft is sent to the opponent, examiner and the supervisor
  - **Should not contain missing parts, i.e this draft shall be complete!**
  - Should be of high quality → This version of the draft is **evaluated by the examiner**

Part 3: Main work, second phase

- Draft is evaluated by the examiner to see
  - How well you have done at this time step (progression)
  - Whether some aspects of the thesis are of poor quality and thus might lead to F unless corrected

- The examiner is asked to mention the critical aspects that might lead to F but it can be hard to do so if there are many problems …
  - Not so good for the grade if the examiner discovers a lot of critical issues …

- The examiner is not supposed to make a detailed review and the feedback on your work but only evaluate your work and possibly give warnings

Revision meeting

- Revision meeting takes 2 hours; it is a closed seminar with the thesis author, opponent, supervisor and examiner

- Review the thesis for which you have been appointed as opponent **in advance:**
  - Go through every piece of the thesis draft in detail and make minor and major comments
  - Everything is subject to criticism, incl. the thesis methodology, results, evaluation metrics
  - Make written list of major issues
  - Mark minor issues (for ex. language errors in the thesis draft)

- Prepare for the Revision meeting for your own thesis
  - Discuss possible opposition time slots with the opponent, examiner and the supervisor
  - Contact your supervisor and request to book a room for meeting at the selected date/time

- Revision meeting is initially led by the opponent; discuss the major issues by going section by section, omit minor errors, approximately 30-45 minutes; send minor errors per e-mail to the thesis author after the meeting

- After the opponent's part, the examiner checks your knowledge of your thesis content and will potentially raise further major issues

Part 4: Finalizing

January 5, 2025	**Submission of the close-to-final thesis to LISAM. *This version will be controlled for the plagiarism.* Send it also to the supervisor, examiner and opponent by email.**
January 12, 2025	Submission of the presentation for the oral defense to LISAM.
January 13, 2025	**Oral defense seminar. In John von Neumann, B building.**

Time frame: revision meeting – final thesis submission

- Make revisions in your own thesis according to what has been taken up by your opponent, examiner and the supervisor
- Prepare your next manuscript version for the defense seminar
- Send your thesis to LISAM, opponent, examiner, supervisor
- Prepare a good presentation for the defense seminar

- Check the revisions made in the thesis on which you are the opponent
- Prepare discussion points for the defense seminar (should take 7-8 minutes)
- Perform on the defense seminar both as speaker and as opponent!

Defense seminar

- Each student presents their thesis for a maximum of 25 minutes
- … followed by the opposition and a general discussion (max. 20 minutes)
  - Audience 5 mins
  - Opponent 7 mins
  - Examiner 7 mins
- After the seminar (or in breaks) each supervisor is recommended to discuss with the student what needs to be adjusted in the manuscript
- **Mandatory to attend all sessions!**

Part 4: Finalizing

January 13, 2025	**Oral defense seminar. In John von Neumann, B building.**
January 20, 2025	Last day to submit final version of the thesis for it to be reported within this semester.
January 27, 2025	Examiner's meeting where the grades are decided and then reported.
March 21, 2025	Last attempt for submitting the final version of the thesis (for students who become delayed; the maximum possible grade is C).
June, 2025	Second opportunity for the oral defense seminar (for students who become delayed).

- **Very short time** to the next submission
  - Minor problems left: revise, submit the final version
  - Major/Many problems left: submit in March
  - Note: ”**last attempt” is only** for students who have gone through all mandatory sessions (incl. defense) but still have some problems left in the report

- If you submit the final manuscript in January, the grade is decided at the examiner meeting
- The course leader (Frank) will tell you your grade
- If you have 120 ECTS, apply for your degree!
  - **It does not happen automatically**

Part 4: Finalizing

How to finalize the work? (if the thesis has grade E or higher):

- Get an ISRN number from Frank after the grades are published
- Use the ISRN and the template provided:
  https://ep.liu.se/publicera-exjobb.aspx
- Consider publishing in E-press:
  https://ep.liu.se/publicera-exjobb.aspx
- Submit your pdf to the administrator Erika Larsson in order to save your PDF into the internal system WexUpp
- You will not get your grade in Ladok until the administrator gets the final PDF from you

Grading

- Grading is based on:
  - **Scientific quality** of the thesis (adequacy and originality of the methods used, significance of obtained results). Note: “originality” is not equivalent to the development of new methods, rather that the methods used have been applied in a novel context.
  - **Technical quality** of the thesis (correctness of language and scientific terminology, implementation, quality of illustrations). This applies to the version of the manuscript submitted before the revision meeting and to the revisions made upon that.
  - **Degree of independence**. The supervisor gives an assessment to the examiner regarding the student’s independence. Listen to the supervisor’s advice but take own initiatives! (but always check with supervisor).

Grading

- Grading is based on:
  - **Progression** (showing sufficient progression between the successive supervision sessions and between different thesis draft versions)
  - **Quality of the performed opposition**, both at the revision meeting and at the final defense
  - **Quality of presentation and communication** (clarity of oral presentation, ability to discuss and defend the thesis, ability to collaborate and communicate with the supervisor)

Use of generative AI

- You have the **full responsibility for your work**, and you must be able to justify the choices you have made in your thesis. Especially:
  - Do not uncritically rely on AI-generated answers since they can be biased or wrong or could just not be adequate for your thesis
    - E.g.: Do not trust summaries of literature – you need to look into cited material yourself! Do not use a lot of text from an AI-tool which is not useful for your thesis!
  - Be aware of the potential for plagiarism and copyright issues since the AI-tool might have generated its answer based on published text or copyright-protected code
  - **Be transparent** with the use of AI-tools: Provide a section in your thesis describing how and to what extent you have used AI-tools in your work
  - Be aware that the examiner will check your knowledge of your text/content at mid-term seminar, revision meeting, and final seminar
  - Discuss with your supervisor if you have questions about the use of generative AI for your thesis

Use of pictures in thesis

- If you use illustations (pictures) from others in your thesis, check if it is allowed to use it
- If allowed, add a reference to the source

Uluru, Australia – by Stuart Edwards, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=1650537

- Example: I need to add a reference to the picture taken from Wikipedia but not to the picture with the horse which I have taken myself

Good luck!