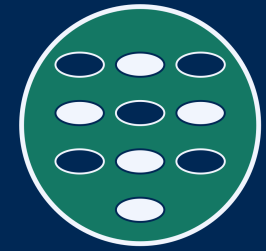


Norges miljø- og
biovitenskapelige
universitet

Institutt for datavitenskap



Digitalisering på Ås

DAT390

Data science seminar

Discussion and feedback on the nearly finished reports
Follow-up on the right to reuse own material
Revised group proposals



Highlight talk schedule for today

Monday, 20th November 2023

Peer feedback from ...

- | | | | |
|---|-------------------------|------------------|--------------------------|
| 15.15 – 15.19 #1 | Rusith C. Hathurusinghe | 15.19 – 15.21 #1 | Maryna Berg |
| Prediction and estimation of indoor radon based on sensory and meteorological data | | | |
| 15.24 – 15.28 #2 | Bastian Undheim Øian | 15.28 – 15.30 #2 | Trishaban Jegatheeswaran |
| Declustering for spatial models | | | |
| 15.33 – 15.37 #3 | Kristoffer Lien | 15.37 – 15.39 #3 | Vegard Molaug |
| Enhancing CRM/ERP data modelling efficiency by MC simulation: Minimum viable data | | | |
| 15.42 – 15.46 #4 | Min Jeong Cheon | 15.46 – 15.48 #4 | Hongpeng Zhang |
| Comparison of survival analysis algorithms on a breast cancer dataset | | | |
| 15.51 – 15.55 #5 | Martin Myklebust | 15.55 – 15.57 #5 | Sushant K. Srivastava |
| (new topic - will discuss how and why it was changed) | | | |



Highlight talk schedule for week 48

Monday, 27th November 2023

Peer feedback from ...

15.15 – 15.19 #1 Petter Bøe Hørtvedt

15.19 – 15.21 #1 Asim Rasheed

Laser disdrometers: Instrument characteristics and uncertainties

15.24 – 15.28 #2 Suhail Rauf

15.28 – 15.30 #2 Bastian Undheim Øian

(undeclared topic)

15.33 – 15.37 #3 Avnik Orbelians

15.37 – 15.39 #3 Mahrin Tasfe

Evaluating digital tools for welfare assessments of Atlantic salmon

15.42 – 15.46 #4 Razieh Kaveh

15.46 – 15.48 #4 Julie Overrein

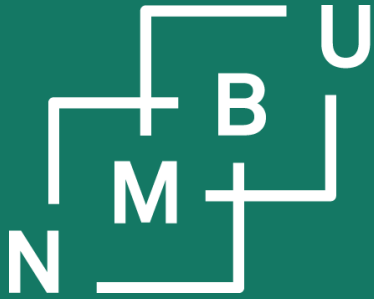
Twitter big data as a source for exoskeleton research

15.51 – 15.55 #5 David Ajaegbu

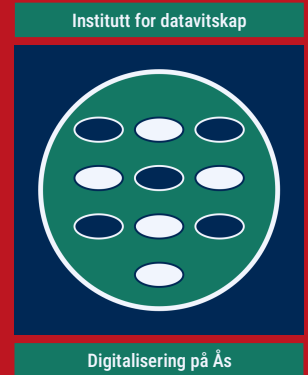
15.55 – 15.57 #5 Awo Arab

Sougata Bhattacharya

Data-driven approach for the prediction of power flexibility



Noregs miljø- og
biovitenskapelige
universitet



Nearly finished reports

Some examples from submitted texts

Take five minutes and think about possible improvements to five examples from some of the nearly finished reports.

Example 2 (this is the very beginning)

Abstract.

With the development of our society, it is important to manage information properly and avoid costly and inefficient systems in place. The better the

Example 3 (text of a short subsection)

In a journal published in Science, the prediction of bus travel times for bus fleets, the information needed for

Example 4 (from an abstract, starting with the very beginning of the abstract)

...ation, Radon is one of the leading causes of lung cancer between 3% to 14% of all lung cancers in a country. ... data science techniques to predict [...] ... to investigate [...]. The ... not have an

Example 1 (from an introduction section)

Getting a new customer can cost 5-12 times more than keeping an existing customer (Singh and Khan, 2023), and a research done by Frederick Reichheld of Bain & Company (the inventor of the Net Promoter score) shows that rise of 5% in customer retention rates can increase profits by 25% (Reichheld, 2001).

Travel prediction for buses

...chers compared different methods for predicting travel times derived from real-time GPS tracking of buses. The results highlighted in the journal were designed to improve bus travel (Fan and ... 5).

Example 5 (from Section 2 of a document)

In 2012, the scholarly work by the authors[5] introduced a groundbreaking Convolutional Neural Network (CNN), one of the largest of its time, applied to a dataset comprising 1.2 million images from the ImageNet ILSVRC-2010 and ILSVRC-2012 competitions.

Self-contained work

Scientific articles, theses, and also the DAT390 report must be **self-contained**.

- An average scientist from your discipline should be able to understand what it says without needing to look up references or other material.
- Define all the mathematical symbols for variables and introduce the variables. If necessary, also introduce notation that may be unfamiliar.
- Introduce all abbreviations, defining them when they are first used.
Exception: Abbreviations that have become part of common language.
 - Only introduce an abbreviation if it is really used again in the text.
 - Check during editing that you don't define an abbreviation twice.

How the DAT390 report will be graded

Scheme for grading DAT390 reports

- 10 Literature review and disciplinary context
- 10 Technical grounding and theoretical insight
- 10 Goal description
- 10 Choice and justification of working methods
- 10 Good practice related to data
- 10 Preliminary results and/or feasibility demonstration
- 10 Analysis and discussion
- 10 Critical reflection and plans or suggestions for future work
- 10 Structure and form
- 10 Language

30%

**Introduction
and theory**

20%

**Methods and
working practice**

30%

**Results and
discussion**

20%

**Form and
academic writing**

A: 90 - 100 points

B: 80 - 89 points

C: 60 - 79 points

D: 50 - 59 points

E: 40 - 49 points

F: 00 - 39 points

Cases with (only) an in-depth literature study

Scheme for grading DAT390 reports

- 10 Literature review and **disciplinary context**
- 10 Technical grounding and theoretical insight
- 10 Goal description
- 10 Choice and justification of **working methods**
- 10 Good practice related to data
- 10 **Preliminary results and/or feasibility demonstration**
- 10 Analysis and discussion
- 10 Critical reflection and plans or suggestions for future work
- 10 Structure and form
- 10 Language

Exception: Literature review only

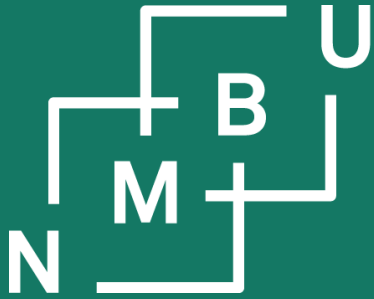
- 10 **Disciplinary context**
- 10 Technical grounding and theoretical insight
- 10 Goal description
- 10 Choice and justification of **literature collection/analysis methods**
- 10 Good practice related to data
- 10 **Systematic literature collection and classification**
- 10 Analysis and discussion
- 10 Critical reflection and plans or suggestions for future work
- 10 Structure and form
- 10 Language

Cases with (only) an in-depth literature study

- This should be done in exceptional cases only.
- Then make sure that you can deliver a strong literature review.
- Obtain a document confirming this and stating that the alternative grading scheme will be applied.
- Include that confirmation in your report as an Appendix.

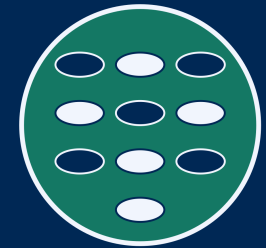
Exception: Literature review only

- 10 Disciplinary context
- 10 Technical grounding and theoretical insight
- 10 Goal description
- 10 Choice and justification of literature collection/analysis methods
- 10 Good practice related to data
- 10 Systematic literature collection and classification
- 10 Analysis and discussion
- 10 Critical reflection and plans or suggestions for future work
- 10 Structure and form
- 10 Language



Noregs miljø- og
biovitenskaplege
universitet

Institutt for datavitenskap



Digitalisering på Ås

Follow-up: Material reuse

Template for agreements with the main advisors

The ideas from last week were mentioned at an institute meeting, and then discussed at a smaller meeting with one of the students participating.

As an outcome of this process, a template is uploaded that can now be used.¹

- 1) Pursuant to NMBU's *retningslinjer «behandling av mistanke om fusk»*, point 3.2h, permission to reuse own material is hereby granted by *aforementioned main advisor* to *aforementioned student* for a master thesis on the thesis topic indicated above.
- 2) The permission extends to the student's own material that has been created more recently than three years before the master thesis due date. It does not extend to older material.
- 3) Content and ideas from the DAT390 report can be reused for the master thesis.
- 4) It is left at the student's discretion whether or how a reference (*i.e.*, citation or other pointer, *e.g.*, a footnote) to the original source of reused own material should be included. An omission to do so will not in itself be judged to constitute academic misconduct.

¹https://home.bawue.de/~horsch/teaching/dat390/material/dat390-template_master-thesis-advisors.docx

Template for agreements with the main advisors

The ideas from last week were mentioned at an institute meeting, and then discussed at a smaller meeting with one of the students participating.

As an outcome of this process, a template is uploaded that can now be used.¹

1) Pursuant to NMBU's *retningslinjer «behandling av mistanke om fusk»*, point 3.2h, permission to reuse own material is hereby granted by *aforementioned main advisor* to *aforementioned student* for a master thesis on the thesis topic indicated above.

2) The permission extends to the student's own material that has been created more recently than three years before the master thesis due date. It does not extend to older material.

3) Content and ideas from the DAT390 report can be reused for the master thesis.

4) It is left at the student's discretion whether or how a reference (*i.e.*, citation or other pointer, *e.g.*, a footnote) to the original source of reused own material should be included. An omission to do so will not in itself be judged to constitute academic misconduct.

¹https://home.bawue.de/~horsch/teaching/dat390/material/dat390-template_master-thesis-advisors.docx

Important: Don't list anybody as co-author

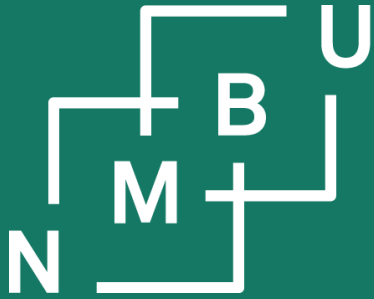
Your DAT390 report must be **your work** – it cannot have a co-author.

- You *can* mention your supervisor(s) as *supervisors*, but not as authors.
- You are not required to mention your supervisor(s).
- If your report has multiple authors, you fail DAT390.

But you collaborated with your supervisor(s) and other students! What now?

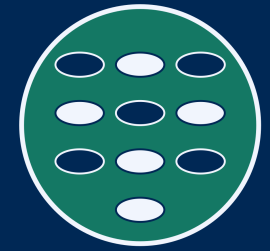
- That is great! We want NMBU to be as strong in research as possible.
- Don't list them as co-authors of your DAT390 report or master thesis.
- **Make anybody else's role explicit through an Appendix.**
 - It is academic misbehaviour if others contributed and you don't tell.
 - You don't need such an Appendix just for your supervisors' role.

Some at REALTEK are experimenting with master theses done in teams of two people. In such exceptional cases, it can be possible to have master theses with multiple authors. That makes it more important to describe who did what.



Noregs miljø- og
biovitenskaplege
universitet

Institutt for datavitenskap



Digitalisering på Ås

Revised group proposals

Revised group proposals

Four groups proceeded to the next round:

- The Original Thesis Titans (#4)
- Visionary Mind (#1)

- Genetic Propheisiers (#5)
- The Forecasters (#8)

These groups were asked to:

- Look into reviewers' feedback and submit a revised proposal.
- Make a very brief statement on what they improved → today.

All groups are now asked to:

- Compare one group against another on one criterion (see Canvas).
- This is to be done as a group task; deadline: End of the month.

Revised group proposals

Four groups proceeded to the next round:

- The Original Thesis Titans (#4): **Revised proposal (link)**
- Visionary Mind (#1)
- Genetic Propheisiers (#5)
- The Forecasters (#8)

These groups were asked to:

- Look into reviewers' feedback and submit a revised proposal.
- Make a very brief statement on what they improved → today.

All groups are now asked to:

- Compare one group against another on one criterion (see Canvas).
- This is to be done as a group task; deadline: End of the month.

Feedback to the *Original Thesis Titans* (#4)

Synergy (4 points)

«sharing ideas, checking each other's work, and helping each other with writing and [...] to support each other emotionally. To make it even better, the plan could include ways to measure how well these ideas are working»

Activity (4 points)

«share ideas, hold brainstorming sessions and [...] a peer review system. [...] the description didn't include [...] how the activities should be done [...]

“**Thesis Pomodoro Plan**” [...] They have done a good job and made it easy to follow. Its rather easy to implement as well, being a **realistic** option.

On whether the activity would succeed involving students beyond the group itself, it would not be that ideal for that.»

Activity (5 points)

«It is realistically executable and seems engaging for students, potentially attracting a wide audience among final year master students in data science.»

Feedback to the *Original Thesis Titans* (#4)

Outcome

- «The proposed activity from the Original Thesis Titans is Thesis Pomodoro Plan.»
- «The group highlights the most essential aspects when considering the outcome of their proposal.»
- «They could be more detailed when looking at [...] when it would start and how many weeks it would be hosted.»
- «Fun and easy activity to be a part of. There are very few “bad” outcomes from this kind of activity.»
- «The group did not mention [...] who will set a date or find a room for everyone to sit in.»
- «This activity can be implemented both physically and virtually, although the group did not discuss this opportunity further.»

Feedback to the *Visionary Mind* (#1)

Synergy (4 points)

«The group identifies one main support mechanism which is [...] group collaboration or open lines of communication/group chat. [...] could have been more creative with ways to support each other and [...] more options.»

Activity (4 points)

«**Data Science Symposium**, [...] an event aimed at bringing together faculty and students [...]. includes student presentations, guest speakers, and panel discussions. It encourages student engagement and provides opportunities for presenting and discussing research. [...] more detail on the execution and management of such a large-scale event could enhance its feasibility.»

Activity (4 points)

«The symposium's objectives, including student presentations, guest speakers, and panel discussions, suggest a well-rounded [...] approach. [...] include [...] details like the date, venue, and registration process [...].»

Feedback to the *Visionary Mind* (#1)

Outcome (4 points)

«The symposium is expected to enhance **presentation skills**, provide exposure to **industry trends**, and facilitate **problem-solving** and **collaboration** among students. The outcomes are clearly aligned with the academic and professional development of students.

However, the proposal could be strengthened by including more specific metrics for measuring these outcomes.»

Outcome (4 points)

«Each part, like "Enhanced presentation skills," "Exposure to Industry Trends," and "Problem Solving and Collaboration," is clearly defined. [...] they might make them even better by brainstorming together for more detailed examples. Talking about communication skills, real-world applications, and working together grabs the reader's attention, showing how everyone benefits. [...] could make it even better by adding more specific examples.»

Feedback to *Genetic Prophesiers* (#5)

Synergy assessment (5 points)

- «emphasizes the **importance of collaboration**»
- «reference the concept of "standing on the shoulders of giants"»
- «**common theme** around the intersection of data science and **bioinformatics** This shared focus can lead to mutual benefits, such as sharing useful articles or datasets.»
- «considering platforms for structured **information sharing**»
- «Regular **group chats** are planned, especially at the start»
- «**diverse strengths** within the group, from numerical expertise to the ability to explain complex ideas. This diversity is seen as a strength»

«I would assign a score of "A" 5 points for Synergy. The group has demonstrated a clear understanding of the importance of synergy and has plans in place to ensure effective collaboration.»



Feedback to *Genetic Prophesiers* (#5)

Activity assessment (4 points)

«activities include:

- Social writing labs, either weekly or bi-weekly [...]
- Outdoor social events [...] to encourage social interaction [...]
- Slack and Discord [...] for sharing resources and [...] communication.
- A GitHub repository for sharing code and research results.»

«proposal is inclusive [...] to the [...] master's student community. [...] especially the outdoor events and the Discord channel are designed to engage [...].

The use of [...] **Slack**, **Discord**, and **GitHub** indicates a strategy tailored to the community's familiarity and comfort, increasing the likelihood of broad participation [...] enhancing both academic collaboration and social interaction.

The proposal realistically addresses the need for both **formal and informal communication and support** during the stressful thesis period»

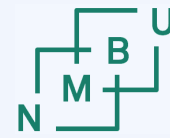
Feedback to *The Forecasters* (#8)

Synergy assessment (3 points)

«the points that are mentioned are more than enough [...]. It would be very fruitful to implement all these points. However, the mentioned points [...] would come to everybody's mind. It will improve the proposal if the last point activities is explained more with added plans for the mentioned activities.»

Synergy assessment (4 points)

«Do they have a clear research question? Yes, their question is: "How can we support each other in our work?" [...] they're communicating regularly, giving feedback, and sharing resources. [...] they plan **weekly meetings, evaluations, a shared literature bank, and social activities**. These plans are doable [...] To improve their collaboration, they could plan for unexpected challenges like conflicts. Having clear conflict resolution methods in place would help. Also, setting up ways to measure their progress, like regular evaluations [...]»



Feedback to *The Forecasters* (#8)

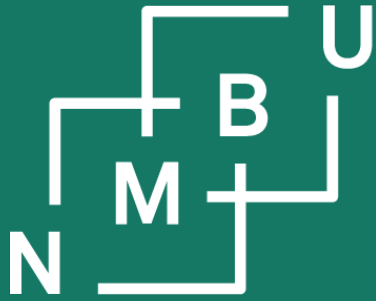
Outcome assessment (4 points)

«Reviewing group no. 8 (The Forecasters) against the "Outcome" criterion.

I think it deserves 4.

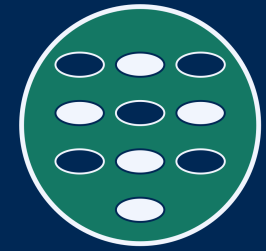
I think they really did a good job. They clearly described how students in the group and the community would benefit from their **Hackthon activity**, both in the report and oral presentation. But [...] maybe they need [...] more "measurable and verifiable" [...] outcomes.

In general, the proposal and presentation was very good. I learned something useful from them.»



Norges miljø- og
biovitenskapelige
universitet

Institutt for datavitenskap



Digitalisering på Ås

DAT390

Data science seminar

Discussion and feedback on the nearly finished reports
Follow-up on the right to reuse own material
Revised group proposals