DAT390
Data science seminar

Observations on the literature review documents

Mid-term evaluation of the course

Form for the mid-term evaluation:

https://nettskjema.no/a/dat390-2024

For best results: Go through the form, two or three people together (5 min.). Then fill it in individually (10 min.).

The form should be anonymous. **Close ▲**

- No information about your user account is automatically stored, even if you happen to be logged in.
- The form should not contain questions that render it possible to identify you based on your response. If the form despite this contains such questions, please contact the person or organization responsible for the form, **martin.thomas.horsch@nmbu.no**.
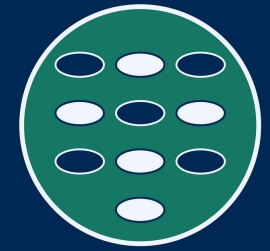
DAT390 2024H mid-term evaluation

What feature, topic, or content from the DAT390 course appears to work best?

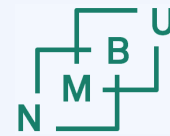What feature, topic, or content from the DAT390 course makes least sense?
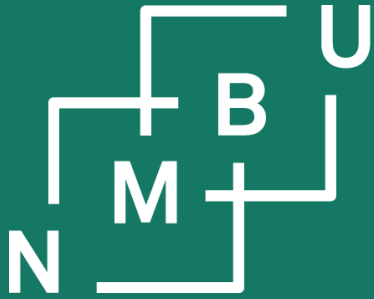
Literature review: Comments

DAT390

21st October 2024

Topic and generality

Your topic must be so specific that you can completely understand the state of the art <u>and</u> make an innovation beyond it, all just within a master thesis.

At the same time, it takes some effort to reflect and zone in on a narrow topic.

Reason to do this **early** on: Reviewing the literature (relevant to your master thesis work) can only be really useful once you have established an exact topic.

Examples for sufficiently narrow topics from our section:

- "Assessing the impact of geopolitical factors and fossil fuel market volatility on electricity price forecasting in European energy markets"
- "Zero/few shot learning for human micro-activity recognition"

But: All of these would be even more suitable as PhD projects (three years!).

It is not necessary that the titles completely show how specific your work is. But a narrow focus must in fact be there, or it is unrealistic in the available time.
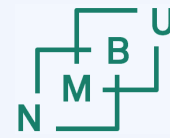
Formatting citations (in the text)

The citations in the manuscript body are annotations, not part of the "real" text. If you imagine reading the text out loud, you should be able to not read the annotation, and the text should still be grammatical.

"As shown in [1], method X works best for Y."

Read: "As shown in, method X works …" ??

"As shown by Authors *et al*. [1], method X works best for Y."
Or: "Method X works best for Y [1]."

No problem here when ignoring the annotation.

When citing in the running text, using parentheses/brackets for the citation, there must be a blank space before the opening parenthesis/bracket.
Like this [1], not like this[1].

If you use superscript citations,[1] they are right at the word or right *after* a punctuation mark.[2] In that case, they are not preceded by a blank space.
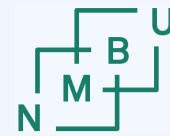
Formatting citations (in the text)

The citations in the manuscript body are annotations, not part of the "real" text. If you imagine reading the text out loud, you should be able to not read the annotation, and the text should still be grammatical.

"As shown in (Authors *et al.*, 2011), method X works best for Y."

Read: "As shown in, method X works …" ??

"As shown by Authors *et al.* (2011), method X works best for Y."

Or: "Method X works best for Y (Authors *et al.*, 2011)."

No problem here when ignoring the annotation.

This is the case even if you use author-year citations.

Then, the part that is in parentheses or brackets is an annotation. The text without the annotation needs to be grammatical.

References / reference list

A few issues related to the cited material and reference lists:

– It is good practice to include a **DOI**, specially as data scientists.
  If you do this, do not include another link/URL in addition to the DOI.

– If you cite an **arXiv** preprint from this year or last year, it is OK.
  – But still check if there is an actual article in the meantime. In
    that case, you must cite the article, not the preprint.
– What if it is an older document on the arXiv?
  – Maybe the result has now been published → cite the article
  – Or it has not been published → that could point to a problem

– You can also cite industry standards or technical white papers from
  governments or associations. But under no circumstances cite the non-
  scientific **"articles" from consultancies**. They are not peer-reviewed and
  there may be an interest behind him that outsiders cannot know.

Reflection on pragmatics

listener

speaker

observer

I don't hold copyright for the material, but am making protected fair use of it.

Reflection on pragmatics

listener
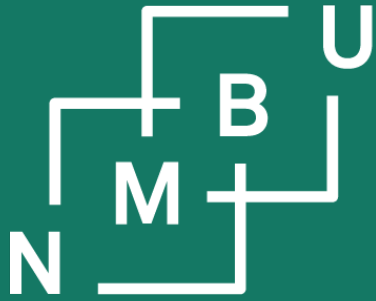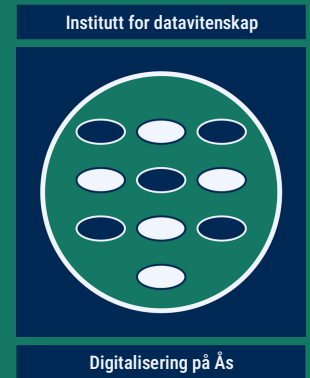
speaker

observer

I don't hold copyright for the material, but am making protected fair use of it.

21st October 2024

DAT390

Data science seminar

Observations on the literature review documents