Academic Writing advice for Machine Learning

For BSc/MSc thesis and papers

Ivo Pascal de Jong, ivo.de.jong@rug.nl

1 Introduction

I find that I regularly give the same writing advice to students. They learn academic writing as a monolith, but expectations vary substantially per (sub)field. This guide gives advice that can apply to writing Machine Learning manuscripts, but will also show some personal preferences. I will try to be clear about what is really expected in Machine Learning writing and what is just my personal preference. Throughout I will point to common mistakes I see, largely influenced by writing advice that is mainly tailored to writing for Psychology or Neuroscience.

2 Tone and Writing Style

In Machine Learning writing we are relatively forward. Top conferences are competitive and reviews are often made with little time. You should be very clear and a bit bold in your writing. Here's some specific pointers:

1. Use active-voice instead of passive voice. Not "a model was trained", but "we trained a model". If you're writing a thesis you can choose to use "I" instead of "We". For a thesis you are the sole author, so "I" makes sense. However, it is a deviation from most of the writing you are familiar with. You can use "we" to make it look more in line with the rest of literature. You will also often see single-author papers use "we".

2. Academic writing should be precise. Try to make sure there's no ambiguity, and use the most-specific phrasing that's available. When you made a model "based on Convolutional Neural Networks", what does "based on" mean? It could mean it is not exactly a CNN, but inspired by CNNs, or it could mean it is just a CNN, or a CNN but with something extra. Be precise.

3. Avoid hyperbole, but state exactly and clearly what your contributions are. You should not say you have made a "breakthrough" or "revolutionised" anything (nor should you say this about the papers you cite), but you can say that you are the first to explore something, proposed a new model that outperforms another by X% or introduced a new framework.

4. Consider that your reader is an expert, but might not be a native English speaker. You do not need to impress your reader by using difficult language, your research will do the impressing. When there is a simpler synonym, please use it. As an example, think about whether there's really any reason to write "utilize" instead of "use". For a deep dive, have a look at this post from ScientistSeesSquirrel.

5. Abbreviations should be used sparingly. You can use it for common abbreviations like Convolutional Neural Networks (CNNs), but you should only introduce at most one novel abbreviation. If your reader doesn't know the abbreviation already you are asking a lot from their working memory. Writing out the words in full actually rarely takes up that much space. All introductions (even common ones) should be written first in full with the abbreviation followed in parentheses.

6. Academic writing should be professional. Avoid slang, colloquialisms, idioms and contractions ("isn't" should be "is not").

7. Using "this/that/it" as pronouns (i.e. to refer to some concept previously introduced) can cause confusion, since it is not always clear what you are referring to. You can attempt to clarify by adding a word to specify ("this model predictions" instead of "this predicts"), or you can avoid the term altogether.

8. Each concept should have one term, and each term should represent one concept. For exam-

ple, what comes out of a Machine Learning model can be called a prediction, an estimate, an inference, an output or a forward pass. If you use those terms interchangeably throughout your paper, a reader might think you are aiming to describe similar but slightly different things (also because in some papers these terms actually do describe different things). Pick one term and use it throughout. This goes directly against general writing advice which says to use synonyms to avoid repetition and make ~~it~~ your work more pleasant to read.

3 Structuring

You are probably familiar with the standard structuring in academic writing: Introduction, Methods, Results, Discussion. This is a very good starting point, but you can (and often should) deviate from this.

When doing multiple experiments, *please* structure your methods and results per experiment. Especially when you do four experiments, if you first describe four methods and then show four results, and then show four times what that means you are asking the reader to remember far too much. In this case it is better to structure it as four experiments, each with methods, results and some interpretation of what those results mean. You can see an example of this in de Jong et al. (2024), where the main body of the paper consists of three experiments each with an introduction, methods, results, discussion and conclusion.

3.1 Sections, Subsections and Paragraphs

The Introduction, Methods, Results and Discussion will typically form the top-level sections. Within these sections LaTeX supports subsections, subsubsections, and paragraphs. Use these headers, but be mindful about them. All section headers should be in Title Case. Most words start with a capital letter, except smaller words like "and", "of", "in", "for".

You should avoid having a subsection header immediately follow a section header. It looks awkward, but it also typically means you are diving into details too soon. A methods section should not immediately start with your model or dataset, but first a high-level about what your methods look like. Only after that should you move into the subsections for your models, datasets, etc..

You do not need to have a sub-sub-section header for thing you describe. For example, if you have a preprocessing pipeline with multiple steps, you are fine to just have a sub-section for Preprocessing, and describe the steps in multiple paragraphs. Each paragraph should describe roughly one thing.

3.2 Introduction

Get to the point of your work very fast. Ideally it should be clear on the first page (but no later than the second page) why I should read your paper/thesis. You do not need to open with the rise of AI and how Machine Learning is increasingly important (actually finding evidence for this is often not easy anyway), but instead move quickly to the gap in the literature. For an extreme example see Suurmeijer et al. (2025).

This means you cannot document the previous literature and existing knowledge in the introduction. To resolve this, you might add a Related Works section after the introduction, as shown in Borszukovszki et al. (2025).

In the introduction you should document what you are researching and why this is relevant. Do not get too hung-up on a Research Question and Hypothesis. If there is no clear hypothesis, you do not need to state it. If the question is an engineering question ("Can we build it?" or "Is my new idea better than the State of the Art?"), that does not make much sense formulating as a proper research question. When there is no nice RQ and hypothesis, instead focus on what the contribution of your work is.

An example of a paper with a (subtle) RQ and without a hypothesis, can be found is Zotos et al. (2025). We knew what we wanted to study, but could see the outcome go either way. At that point, it does not make sense to create an awkward hypothesis. In Manivannan et al. (2024) we did have a very clear hypothesis. The study was motivated by a theory known from literature that things should work in a certain way, and we wanted to validate whether this theory worked in practice (it did not!).

3.3 Methods

I find that students are typically good at writing the methods section. Make sure you clearly document what you did, and why you did it this way. For any non-trivial decisions, you should provide argumentation for why you do it this way. You should only justify why your way is the right way. You do not need to describe why there might be problems with your methods, that comes later in a limitations section.

The requirement for what constitutes argumentation is often lower than what students expect. It can be following existing literature, easy to reproduce, computationally affordable, or intuitive. Almost always you did have a reason for your choice, and just writing it down can be sufficient.

However, in writing the papers the bounds of your own skills and timelines are typically not a valid argument. Arguments like simplicity (Occam's razor) and "scope" (aim) of the study typically are a good substitute. If your aim is to show a minimum working example or do a feasibility study, then doing your experiments with small toy datasets is perfectly acceptable.

Some people like to end the methods section with describing the metrics and how things will be evaluated. This is good, but avoid requiring the reader to keep many details in their working memory before moving to the results section. Do not describe what the plots will look like, but you may describe abstractly which analyses will be done or which aspects will be investigated.

3.4 Results

Typically your results section revolves around Figures and Tables. Like other sections, introduce the results section, and then move first to the most important result. After that you can proceed to the additional analyses or experiments.

Results are best written by taking a Figure / Table and explaining what is in that Figure / Table. You can pretend the reader is blind, so you have to point out to them what is to be seen in the Figure / Table. You can already interpret what this means. Do not present the results neutrally as describing exactly what they are, but point the reader to the interesting parts and tell them what that means. If you feel like interpretation belongs in the Discussion section: In the results section you give a low-level interpretation of each finding. In the Discussion this will become more high-level and you will then look at the bigger picture. You can see this distinction fairly clearly in Zotos et al. (2025).

In the results section you do not need to present all the results that you generated. Instead, you have taken the time to interpret the results, and you present the results that show evidence related to your interpretation. This is much more effective for conveying your research than the "objective" ideal of presenting all results. To prevent cherry-picking and to show the extra work that you did, you can put all of the other results in the Appendix.

3.4.1 Statistical Testing

Statistical testing is overrated in Psychology, but underrated in Machine Learning. As a result, for Machine Learning papers statistical testing is not expected and you are not required to deliver statistical tests.

However, you *are* required to deliver a measure of variance, so we can establish intuitively whether differences are likely due to chance or not*. What exactly you should be measuring variance over depends on your experiment, but often running your experiments with different seeds is a valid option.

When I say statistical testing is underrated, I mean that if you see a good opportunity to do statistical testing in a way that makes sense, I think you should (even if the Machine Learning community would not expect you to). Things like running a (paired) t-test or ANOVA are often easy and can help solidy whether something is a coincidence, or Pearson/Spearman for checking for correlations. `https://apastyle.apa.org/instructional-aids/numbers-statistics-guide.pdf`APA has a guide that is generally good, but I personally prefer p-values in scientific notation than truncated to give slightly more precision $p = 3.4 \times 10^{-10}$ instead of $p < 0.001$

Whenever you report multiple p-values you should do a correction for multiple comparisons. Generally Bonferroni correction is simple and easy. You can either do this by scaling the p-values or scaling the significance threshold α based on the

*Please specify whether you are showing variance, standard deviation, confidence intervals, ...

number of tests. When you find that you are doing far too many statistical tests and Bonferroni correction is too conservative then it might make more sense to not report them.

If you have many statistical tests to present, please do not present them in the "proper" written way, but use a table. Also make sure you are conscious about how many decimal points you use. Often using too many makes it harder to read and does not tell the reader anything more.

3.5 Discussion

Open the discussion with the main takeaway from your results. In summary, what does it all say? Since this should be higher-level, you want to avoid the specific metrics and details and instead talk about what you have shown.

After that comes the *real* discussion where you describe how your results relate to other literature, your limitations and suggestions future work. Please avoid trivial limitations and future work suggestions such as bigger models, larger datasets and more participants which apply to almost every paper, but try to find something insightful and unique about your paper. What might be the real holes in your study that could partially invalidate it? What are new research questions that spawn from your project? Ideally, from the suggestions of future work a reader should be able to sketch a new study that you would indeed find interesting, promising and will give novel findings.

To end on a high-note, you want to close by concluding positively (you do not want to end on everything that is wrong with your research). You can do this as a separate Conclusion section. In the conclusion make sure you very clearly show what the contributions are. One way to do this is with an itemized list of things you contributed, see Borszukovszki et al. (2025) for an example. What you make very clear what you found and contributed, a reader, reviewer or grader will know exactly what your claim is and can validate whether that is supported by your evidence.

Avoid concluding that your study *underscores the importance* of something. That is not a claim that can be (in)validated from your results, and does not exactly have a meaning.

4 Figures and Tables

Figures and Tables also have specific requirements that seem to be often broken.

Firstly, they should always float to the top (preferred) or bottom of the page, and appear close to where they are referenced in the text. They should also appear in the order they are referenced. To make sure this happens, include them as `\begin{figure}[t]` (top) or `\begin{figure}[b]` (bottom), and play around with the sequence relative to the text until things add up. You can also use `\FloatBarrier` from `\usepackage{placeins}`. However, getting this exactly right is the last thing you should do before submitting the paper. Often when you add paragraphs and change text things start to move again. Also make sure the Figure do not cause awkward gaps of whitespace.

Second, they should have an informative caption. The caption should describe what is to be seen, but also a quick takeaway from the image. It is not unusual to see captions with 2-3 lines. As an example, try reading Mucsányi et al. (2024) by only reading the captions of the Figures. This is a great way to get the gist of the whole paper. Because your Figures have a caption, *they do not need to have a title in the image.* Those titles are helpful when handling images on their own, but make less sense when there's also a caption below it. If you want to be very fancy with it, have a look at Borszukovszki et al. (2025) for how we embedded the legend in the caption.

Third, they should be of high quality. Plots should always be imported as vector graphics. The easiest way is to save your `matplotlib` figures with `plt.save('some_figure.pdf')`. This way your plots will have infinite-zoom, and will never look pixelated. They should also be (mostly) legible if printed. This typically means you want to increase the font-size in matplotlib, or decrease the `figsize`, which will also make all the lines thicker. If you need to space, you can also make your figure span both columns.

For Tables there is a rule that professional Tables should not have vertical lines. Use the `booktabs` package with `\toprule`, `\midrule` and `\bottomrule`. Tables 4.1 and 4.2 show the difference. To help the reader, you should also highlight the things they should look at. This is typ-

Ugly Table	Column1	Column2
Row 1	0	**1**
Tow 2	**1**	0

Table 4.2: This is a nice table useing booktabs. The caption can be either at the top or the bottom, depending on the template. Bold indicates the largest numbers.

ically done by making the best-performing method in bold. This should also be mentioned in the caption (even though it's often obvious). You can see a more elaborate example of this in Zotos et al. (2025).

Ugly Table	Column1	Column2
Row 1	0	0
Tow 2	0	0

Table 4.1: This is an ugly table. It should not be used.

There are typically two special Figures in every Machine Learning paper that are crucial, which I will describe below.

4.1 Figure 1.

The Figure 1. in Machine Learning papers is a Figure (or Table) on the first (maybe the second) page that immediately shows what this paper is about and why it is great. The Figure 1 from Mucsányi et al. (2024) became the thing people remember about that paper.

Sometimes you will need to be a little creative to make this happen. Visual examples are often helpful, but illustrations and diagrams can also work. In Suurmeijer et al. (2025) we generated images that are purely illustrative, but that do explain the point of the paper. This is an excellent way to draw attention and show why your paper is worth reading. It can also be what you later use to post what you did on LinkedIn.

4.2 Figure 2.

The other crucial figure is the one that supports your methods section (this is not necessarily Figure 2, but can be some other number). Often methods sections get complicated and have many moving parts. Creating a diagram in draw.io or powerpoint can allow you to create a nice and clear visual explanation of your methods.

When running experiments with humans, typically this can show the stimuli or the series of experimental conditions.

5 Math

Other academic writing guides often ignore math, but this is often important for Machine Learning. Each variable ν in an equation should be defined. To help readers keep track, it helps to add the term of what it means. If you define a temperature T do not say "As T increases ..." but instead write "As the temperature T increases ...".

Equations should be numbered and follow regular punctuation. Let us define some equation for $\hat{y}$ as

$$\hat{y} = \frac{\exp(-d_i^2/T)}{\sum_j \exp(-d_j^2/T)}. \tag{5.1}$$

We can see it is followed by a period, and it follows naturally as reading. In this case we need to also still define a distance d_i to make the equation complete.

If you use multiple equations from different sources, keep track that each variable is defined exactly once. I can not define a time T for my next equation, because T is already a temperature.

There is a complicated set of conflicting conventions. For examples, a capital letter indicates a matrix, or a Random Variable, but here temperature T is a scalar that is capitalized by convention (there's a long history that goes back to the physics of heat treating metals to reduce internal stresses). It is generally good to read other related literature and see which conventions they follow, and combine this with knowledge that you have about standard math notation.

6 Doing the writing

Probably the hardest part of writing is actually sitting down and doing the writing. It is simply a hard thing to do, but it does get better with practice. I will give some advice that you might try if you find that it is difficult to get the words out:

1. Ignore all ideas about what is good writing. That advice is actually only for editing. Just accept that whatever comes out is what comes out, even if it is not formulated the way you want or does not sound right. When editing you can fix those things, but when writing a first draft your work does not need to be good.

2. Buy a rubber duck and explain the situation that you want to write to the rubber duck. Then, write down the explanation you gave to the duck. You can substitute the rubber duck for a human being and you might even use a transcription AI to do the writing for you. I have never tried these transcription tools, but I have regularly subjected my dog to hearing me talk about my next paragraph,

3. Go for a short 5 minute walk. During the walk you can think about what you actually want to say. Movement is good for the brain, which is good for ideas.

4. Do not use LLMs to write anything you care about. The difficult part about writing is generating new and interesting ideas and arguments. LLMs can generate ideas and arguments, but they will not be new.

5. Set a deadline for completing a (sub)section. Say you will write Section 4.2 today before 17:00. Towards the end of the day you will rush yourself, take short-cuts and write a low-quality version, but it will be done. Then you can fix things in editing.

6. Get comfortable being uncomfortable. Assign yourself a timeframe in which you will do the writing. Do not wait until you feel motivation or have ideas. Instead, spin around in your chair bored until the time is over. This will actually push you to do the writing. Make sure your phone is not in the same room as you.

7. Write (and give) a presentation first, then the paper/thesis. This way you figure out the whole narrative and all the topics that need to be covered, so all that is left to do is put the words on paper.

7 Citing

Make sure you use citations generously. When possible, you should cite for: claims, methods, models, metrics, equations, related topics, helpful review papers, architectures, (open source) libraries, datasets, and probably more. When citing papers for a named contribution, it is helpful to put the citation immediately after the name. For example, the concept of Disentanglement Error (de Jong et al., 2024) should be cited like this. This style makes it easy to see the source of each item.

Towards the end, you should also check your references to see whether they are correct. A good starting point is usually the bibtex you find in Google Scholar, but afterwards check whether the styling is consistent and check there are no mistakes. Google Scholar automatically generated the bibtex, and authors and journals are not able to fix them. In this case, you can see that Zotos et al. (2025) has “Nlp methods” which should probably be “NLP methods”. Many journals also include bibtex, which is often more correct.

ArXiv papers are pre-prints, which indicates it has not yet been peer reviewed. To some readers, this indicates a lesser status. When citing papers as arXiv preprint, please check whether the paper might also already be published somewhere. Then it is better to cite the published version (unless there is something specific about the arXiv version that is not in the published version, but that is very rare). You can find the published version by clicking “All X versions” on Google Scholar.

I typically do not worry too much about the specific styling, as long as it is consistent. The LaTeX template you use will probably handle that for you.

8 Appendices

Please, use appendices freely. You can insert full-size figures there, copies of results under slightly different conditions with similar outcomes, and many qualitative examples. Appendices typically look nicest full-width.

It is also perfectly allowed to write text in Appendices. If you want, you can add whole experiments in the Appendix where you describe that method, and those results in an Appendix.

To make sure people actually look at your appendix, it is often good to refer to it from within the text. You can say "As we show in Appendix Q, ...".

9 Specifics for MSc thesis

For the AI/CCS MSc thesis at RuG there is no template. You are free to style it how you see fit (within sensible bounds). A very helpful rogue student (Manvi Agarwal) made their own template and posted it on Overleaf `https://www.overleaf.com/latex/templates/master-thesis-template-university-of-groningen/txxpkkjmfksx`. If you are not sure what styling to use, use this as a starting point, and possibly as your ending point.

The MSc thesis often has a very extensive theoretical background, much longer than you see in papers. In this, you can showcase all of the things that you've learned about your project. It does not need to include strictly relevant information. It serves as a useful introduction into the broader topic that you study.

The MSc thesis typically also has an acknowledgements section. It is customary to thank your supervisors here in some capacity (withholding this would be a clear statement that your supervisor was terrible!), but more importantly you can use this space to also thank friends, family, colleagues, pets, and a developer of some Github repository who helped you with debugging while you were trying to use their tool.

10 Additional sources

If you want even more advice about academic writing, I recommend this blog post on getting an ML paper accepted: `https://maxwellforbes.com/posts/how-to-get-a-paper-accepted`. It focuses on writing for peer-reviewed top conferences, which is very different from writing a thesis, but some things are the same. In both cases you have someone evaluating the quality of your work with human biases. If you can show your reviewer/examiner that your work is interesting and a meaningful contribution it can help you get accepted/get a high grade.

Appendix A from the Neural Networks projects instructions (`https://www.ai.rug.nl/minds/uploads/SemesterProjectInstruction_NN.pdf`) gives advice on writing technical writing. Many of these things also apply to your thesis.

References

Borszukovszki, M., De Jong, I. P., & Valdenegro-Toro, M. (2025). Know what you do not know: Verbalized uncertainty estimation robustness on corrupted images in vision-language models. *arXiv preprint arXiv:2504.03440*.

de Jong, I. P., Sburlea, A. I., & Valdenegro-Toro, M. (2024). How disentangled are your classification uncertainties? *arXiv preprint arXiv:2408.12175*.

Manivannan, P., de Jong, I. P., Valdenegro-Toro, M., & Sburlea, A. I. (2024). Uncertainty quantification for cross-subject motor imagery classification. *arXiv preprint arXiv:2403.09228*.

Mucsányi, B., Kirchhof, M., & Oh, S. J. (2024). Benchmarking uncertainty disentanglement: Specialized uncertainties for specialized tasks. *Advances in neural information processing systems*, *37*, 50972–51038.

Suurmeijer, J., de Jong, I. P., Valdenegro-Toro, M., & Sburlea, A. I. (2025). Uncertainty quantification for motor imagery bci–machine learning vs. deep learning. *arXiv preprint arXiv:2507.07511*.

Zotos, L., de Jong, I. P., Valdenegro-Toro, M., Sburlea, A. I., Nissim, M., & van Rijn, H. (2025). Nlp methods may actually be better than professors at estimating question difficulty. *arXiv preprint arXiv:2508.03294*.