Seminal ML paper summary

Learning objectives

With this task, you will:

- Read and understand in great detail a very impactful ML paper
- Practice your writing skills
- Learn to summarize large volumes of relatively complex text in a few sentences
- Learn to write a real conference paper
- Get experience of the responsible use of ChatGPT and similar tools
- Force the instructors to read all these impactful papers by reading your summaries :)

What should be done, how, and when?

Choose and read a paper (**1 out of 12 proposed at the bottom of the document**) and write a concise summary essay based on this paper.

A summary essay is a concise but informative distillation of the main points made in the original piece. It would be hard to describe the process better than it was done here or here (or you can ask ChatGPT for instructions). Please, follow the aforementioned guidelines, and below we present some organizational do's and don'ts.

Structure

Your summary should be a **self-sufficient research paper** itself. Therefore, we want you to structure it as a tiny conference paper. We propose that you work in **Overleaf** and use the ICML paper format, here is one of the past conference templates (they rarely change from year to year). If you particularly don't like it, you can take a template from another conference, but make sure it is in **two-column format**. Your paper should have:

- **Title** (should state that your work is a summary)
- **Author** (you) with an affiliation (your university, institute, …)
- **Abstract**
- **Introduction**
- Main body parts: **Related work**, **Methods**, **Results**, **Discussion**, etc. You don't need to have all of these, since it is a short summary and your sections most likely will be merged, e.g., your related work will likely go into introduction. However, the logic should still be clear and the paper should be concise. (As for another typical mistake, don't use headings like "Main" or "Main body").
- **Conclusion**
- **References** (at least 3-5 relevant references cited in your text: original paper, background, related work, etc.)
- Optionally, you can write **Acknowledgement** and/or add an **Appendix** (or several).

Every part of the conference paper has its own purpose and historical background. If you are unsure what the difference between all these parts is, you can understand it from the paper you are summarizing or check some tutorials, like this one.

The paper should be **no longer than two pages**, excluding the references, acknowledgments, and appendices. In other words, your conclusion should end before the end of the second page. The paper can be shorter, but not too short, and you'll likely lose a lot of information.

NB: Your paper should be clearly positioned as a summary, and you should **not** pass off the authors' contributions as your own. You can use constructions like 'The authors proposed a new method…', or 'Vaswani *et. al.* showed that…'.

Formatting

The formatting of your paper should also meet the requirements of a conference paper. You should use proper citing, illustrations, references, tables, algorithms, etc. Here we enumerate the most important rules:

- Your figures, tables, formulae, algorithms, etc. should be properly **captioned** and **enumerated**. You cannot simply add a screenshot of the caption from the paper. There should be no unreferenced content, so make sure that everything you add is properly **mentioned in the text**. If you use the template we proposed, it will be very convenient to organize everything nicely.
- If you use tables or formulae, they should be done in LaTeX + Markdown, also not screenshots.
- Citation is always **part of the sentence**, not outside of it [2]. And usually, sentences must be readable without citations as well. Not like: "In [1] a new method was proposed", but rather "A new method was proposed [1], that ...". Normally, there are no citations in the abstract.
- The graphic/tabular content you take from the original paper should always be cited.
- The list of references should correspond to the citations and be properly formatted. Once again, if you use the Overleaf template, it will be very easy to organize it, and a very important skill to master.
- Don't use the I pronoun. In scientific texts, it's either **we** (even if there is only one author), or **passive voice**. Once again, "we" in this case is you, and you are summarizing, the authors of the original paper are **not** the "we" in your summary.

If you are not sure if you are using the right formatting, you can always check the conference template we have given you. There are **examples** of formatting of everything you need and more, and the rules for the submission to be accepted. In a real conference, if a paper doesn't meet the formatting rules, it gets a **desk-reject**.

NB: Submit only a **PDF** file, **other formats will not be accepted**. Overleaf can compile your paper into a PDF and you can download it.

Originality and ChatGPT

Even though you are summarizing an existing work, the content of your summary should be a **new text**. Therefore, we will additionally check your work for **plagiarism**. You should not have any unedited extracts from the original paper or its other summaries (you can use any plagiarism detection software yourself, e.g., urkund). Of course, this does not apply to specific terms or acronyms from the paper. Your originality should not harm the correctness. You can always **cite** the original paper when you want to draw attention to how a certain thing is proposed to be done by the authors.

You will likely be tempted to use ChatGPT or other generative tools to help you with the writing. We will **not** advise against that, but we would like to **warn** you about the following:

- LLMs **hallucinate**. They tend to invent papers, authors, and terms or make errors in formulae, algorithms, code, calculations etc., due to autoregressive generation.
- Generative models can mix up concepts and were not trained on the most recent data.
- Free versions won't provide you with the whole summary, so your text can lose in logic, structure, style, etc.
- This one is unlikely, but the generative models can produce text that overlaps the original paper.

Regarding all this, if you are using generative models, you are responsible for preventing these risks. We will check your text for logical errors, misconceptions, and made-up content.

Grading

We have five criteria, each of which gives a part of the score:
1. **Content** (5 pt.) – everything related to how the summary corresponds to the actual paper – does it have all the necessary information, does it have hallucinations or irrelevant facts, etc
2. **Structure** (2.5 pt.) – looks like a paper, has sufficient structured sections, is of the required length, etc.
3. **Formatting** (2.5 pt.) – everything related to figures, tables, references, citations, etc.
4. **Plagiarism** (2.5 pt.) – deductions if the summary has unreferenced extracts (e.g., one sentence) from the original paper/other papers/other summaries/guides, etc.
5. **Style and Grammar** (2.5 pt.) – is the text readable, does it have grammatical errors, or is it uncommon for research papers language

List of papers:

Try not to use article length as the main criterion for decision-making. Each one of these papers is precious. Let your heart decide :)

01 - Attention Is All You Need (2017)
https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

02 - XGBoost: A Scalable Tree Boosting System (2016)
https://arxiv.org/pdf/1603.02754.pdf

03- U-Net: Convolutional Networks for Biomedical Image Segmentation (2015)
https://arxiv.org/pdf/1505.04597.pdf

04 - Deep Residual Learning for Image Recognition (2015)
https://arxiv.org/abs/1512.03385

05 - Dropout: A Simple Way to Prevent Neural Networks from Overfitting (2015)
http://jmlr.org/papers/volume15/srivastava14a/srivastava14a.pdf

06 - Batch Normalization: Accelerating Deep Network Training by Reducing
Internal Covariate Shift (2015)
http://proceedings.mlr.press/v37/ioffe15.pdf

07 - Generative Adversarial Nets (2014)
http://datascienceassn.org/sites/default/files/Generative%20Adversarial%20Nets.pdf

08 - Playing Atari with Deep Reinforcement Learning (2013)
https://www.cs.toronto.edu/~vmnih/docs/dqn.pdf

09 - Visualizing and Understanding Convolutional Networks (2013)
https://arxiv.org/pdf/1311.2901.pdf

10 - Efficient Estimation of Word Representations in Vector Space (word2vec, 2013)
https://arxiv.org/pdf/1301.3781.pdf

11 - ImageNet Classification with Deep Convolutional Neural Networks (AlexNet paper, 2012)
https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

12 - Random forests (2001)
https://www.stat.berkeley.edu/~breiman/randomforest2001.pdf

BTW, if you feel that some paper is missing from this list, let us know.