

NeurIPS Paper Checklist Guidelines

The NeurIPS Paper Checklist is designed to encourage best practices for responsible machine learning research, addressing issues of reproducibility, transparency, research ethics, and societal impact. The checklist is included in [the LaTeX style file](#). Do not remove the checklist: **The papers not including the checklist will be desk rejected.**

All submissions must be in PDF format, and in a single PDF file include, in this order, (1) the submitted paper; (2) optional technical appendices that support the paper with additional proofs, derivations, or results; (3) the NeurIPS paper checklist. The checklist should therefore follow the submissions's references and the (optional) supplemental material. The checklist does NOT count towards the page limit.

Please read the checklist guidelines carefully for information on how to answer these questions. For each question in the checklist:

- You should answer yes, no, or n/a.
- NA means either that the question is Not Applicable for that particular paper or the relevant information is Not Available.
- If you would like to explain your answer (e.g., explain why the answer is "no", or provide more detail reviewers), consider referencing the section(s) of the paper that provide support for your answer.
- You can also optionally write a short (1–2 sentence) justification of your answer.

The checklist answers are an integral part of your paper submission.} They are visible to the reviewers, area chairs, senior area chairs, and ethics reviewers. You will be asked to also include it (after eventual revisions) with the final version of your paper, and its final version will be published with the paper. **The checklist answers are visible to the reviewers, area chairs, senior area chairs, and ethics reviewers.**

Reviewers will be asked to use the checklist as one of the factors in their evaluation. While "yes" is generally preferable to "no", it is perfectly acceptable to answer "no" provided a proper justification is given (e.g., "error bars are not reported because it would be too computationally expensive" or "we were unable to find the license for the dataset we used"). **In general, answering "no" or "n/a" is not grounds for rejection.** While the questions are phrased in a binary way, we acknowledge that the true answer is often more nuanced, so please just use your best judgement and write a justification to elaborate. All supporting evidence can appear either in the **main paper or the supplemental material provided in the appendix**. If you answer "yes" to a question, in the justification please point to the section(s) where related material for the question can be found.

We provide guidance on how to answer each question below. You may additionally refer to the [blog post](#) from 2021 introducing the checklist to learn more about its motivation and how it was created.

1. **Claims:** Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? Claims in the paper should match theoretical and experimental results in terms of how much the results can be expected to generalize. The paper's contributions should be clearly stated in the abstract and introduction, along with any important

assumptions and limitations. It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper. Enter yes, no, or an explanation if appropriate. Answers are visible to reviewers. Claims in the paper should match theoretical and experimental results in terms of how much the results can be expected to generalize.

- The paper's contributions should be clearly stated in the abstract and introduction, along with any important assumptions and limitations. It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations:** The authors are encouraged to create a separate "Limitations" section in their paper. The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally).

- The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
- The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
- The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
- We understand that authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection. It is worth keeping in mind that a worse outcome might be if reviewers discover limitations that aren't acknowledged in the paper. In general, we advise authors to use their best judgement and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.

3. **Theory, Assumptions and Proofs:** If you are including theoretical results, did you state the full set of assumptions of all theoretical results, and did you include complete proofs of all theoretical results? All assumptions should be clearly stated or referenced in the statement of any theorems. The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, authors are encouraged to provide a short proof sketch to provide intuition. Enter yes, no, n/a, or an explanation if appropriate. Answers are visible to reviewers.

- Did you state the full set of **assumptions** of all theoretical results? All assumptions should be clearly stated or referenced in the statement of any theorems.
- Did you include complete **proofs** of all theoretical results? The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, authors are encouraged to provide a short proof sketch to provide intuition. You are encouraged to discuss the relationship between your results and related results in the literature.

- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. Experimental Result Reproducibility: If the contribution is a dataset or model, what steps did you take to make your results reproducible or verifiable? Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general, releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), release of a model checkpoint, or other means that are appropriate to your research. Enter yes, no, n/a, or an explanation if appropriate. Answers are visible to reviewers.

- While NeurIPS does not require releasing code, we do require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example...
 - If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
 - If the contribution is primarily a new model architecture, the paper should describe the architecture fully.
 - If the contribution is a new model (e.g., a large language model), then there should either be a way to **access** this model for reproducing the results or a way to **reproduce** the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
 - We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. Open Access to Data and Code: If you ran experiments, did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? Please see the NeurIPS code and data submission guidelines for more details. While we encourage release of code and data, we understand that this might not be possible, so no is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark). At submission time, to preserve anonymity, remember to release anonymized versions. Enter yes, no, n/a, or an explanation if appropriate. Answers are visible to reviewers.

- Did you include the code, data, and instructions needed to **reproduce** the main experimental results (either in the supplemental material or as a URL)?
- The instructions should contain the exact command and environment needed to run to reproduce the results.
- Please see the NeurIPS [code and data submission guidelines](#) for more details.

- Main experimental results include your new method and baselines. You should try to capture as many of the minor experiments in the paper as possible. If a subset of experiments are reproducible, you should state which ones are.
- While we encourage release of code and data, we understand that this might not be possible, so "no because the code is proprietary" is an acceptable answer.
- At submission time, to preserve anonymity, remember to release anonymized versions.

6. Experimental Setting/ Details: If you ran experiments, did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? The full details can be provided with the code, but the important details should be in the main paper, and information about how hyperparameters were selected should appear either in the paper or supplementary materials. Enter yes, no, n/a, or an explanation if appropriate. Answers are visible to reviewers.

- The full details can be provided with the code, in appendix or as a supplement, but the important details should be in the main paper.
- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. Experiment Statistical Significance: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. Experiments Compute Resource: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

- The answer NA means that the paper does not include experiments.
- The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
- The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute. The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper)

9. Code Of Ethics: Have you read the [NeurIPS Code of Ethics](#) and ensured that your research conforms to it? Enter yes, no, or an explanation if appropriate. Answers are visible to reviewers.

- If you have special circumstances that require some sort of deviation from the Code of Ethics, please explain it here. Please make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in your jurisdiction).

10. Broader Impacts: If appropriate for the scope and focus of your paper, did you discuss potential negative societal impacts of your work? Please see the Paper Checklist Guidelines for detailed instructions and examples of points that you may choose to discuss. Enter yes, no, n/a, or an explanation if appropriate. Answers are visible to reviewers.

- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- We expect many papers to be foundational research and not tied to particular applications, let alone deployments. However, if you see a direct path to any negative applications, you should point it out. For example, if you improve the quality of generative models, you might point out that your approach can be used to generate Deepfakes for disinformation. On the other hand, if you develop a generic algorithm for optimizing neural networks, you do not need to mention that this could enable people to train models that generate Deepfakes faster.
- Consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, you could also discuss any mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. Safeguards: Do you have safeguards in place for responsible release of models with a high risk for misuse (e.g., pretrained language models)? Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model. Enter yes, no, n/a, or an explanation if appropriate. Answers are visible to reviewers.

- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort

12. **Licenses:** If you are using existing assets (e.g., code, data, models), did you cite the creators and respect the license and terms of use? Cite the original paper that produced the code package or dataset. If possible, include a URL. Be sure to check the original license and respect its conditions. Enter yes, no, n/a, or an explanation if appropriate. Answers are visible to reviewers.

- Cite the original paper that produced the code package or dataset.
- Remember to state which version of the asset you're using.
- If possible, include a URL.
- State the name of the license (e.g., CC-BY 4.0) for each asset.
- If you scraped data from a particular source (e.g., website), you should state the copyright and terms of service of that source.
- If you are releasing assets, you should include a license, copyright information, and terms of use in the package. If you are using a popular dataset, please check paperswithcode.com/datasets, which has curated licenses for some datasets. You are also encouraged to use their [licensing guide](#) to help determine the license of a dataset.
- If you are repackaging an existing dataset, you should state the original license as well as the one for the derived asset (if it has changed).
- If you cannot find this information online, you are encouraged to reach out to the asset's creators.

13. **Assets:** If you are releasing new assets, did you document them and provide these details alongside the assets? Researchers should communicate the details of the dataset or the model as part of their submissions via structured templates. This includes details about training, license, limitations, etc. Enter yes, no, n/a, or an explanation if appropriate. Answers are visible to reviewers.

- The answer NA means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and Research with Human Subjects:** If you used crowdsourcing or conducted research with human subjects, did you include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)? Including this information in the supplemental material is fine, but if the main contribution of your paper involves human subjects, then we strongly encourage you to include as much detail as possible in the main paper. According to the NeurIPS Code of Ethics, you must pay workers involved in data collection, curation, or other labor at least the minimum wage in your country. Enter yes, no, n/a, or an explanation if appropriate. Answers are visible to reviewers.

- Did you include the full text of instructions given to participants and screenshots, if applicable? Including this information in the supplemental material is fine, but if the main contribution of your paper involves human subjects, then we strongly encourage you to include as much detail as possible in the main paper.

15. **IRB Approvals:** Did you describe any potential participant risks and obtain Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your institution), if applicable? Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper. For initial submissions, do not include any information that would break anonymity, such as the institution conducting the review. Enter yes, no, n/a, or an explanation if appropriate. Answers are visible to reviewers.

- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper. We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution. For initial submissions, do not include any information that would break anonymity, such as the institution conducting the review.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage:** Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

- The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our [MainTrackHandbook](#) for what should or should not be described.