

[Login](#)

Reproduction Guidelines

These guidelines are based on [NeurIPS 2021 Paper Checklist](#). Minor modifications to this checklist were done by Razvan Pascanu.

The CoLLAs guidelines are meant to encourage best practices for responsible machine learning, addressing issues of reproducibility, transparency, research ethics and social impact. Please read the following questions and reflect on what the answer would be for your current submission. You are encouraged to take a look at the checklist early on as it likely will positively influence the way you do your research and write the paper, rather than treating it as an afterthought.

Reviewers can rely on this list to motivate their evaluation and argue for lack of transparency or reproducibility. Therefore the list can act to help the reviewers as well to structure their feedback.

We also urge the reader to look at the following [blogpost](#) from NeurIPS 2021 Program Chairs introducing and motivating the list of questions that we relied on and used here.

Please read the following questions and reflect on what would be the answer for your work:

- 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.
- Have you read CoLLAs code of ethics and ensured that your paper conforms to them? Please read CoLLAs code of ethics.
- Did you discuss any potential negative societal impacts of your work?
 - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), environmental impact (e.g., training huge models), fairness considerations (e.g., deployment of technologies that could further disadvantage historically disadvantaged groups), privacy considerations (e.g., a paper on model/data stealing), and security considerations (e.g., adversarial attacks).
 - We expect many papers to be foundational research and not tied to particular applications, let alone deployments, but being foundational does not imply that research has no societal impacts. If you see a direct path to any negative applications, you should point it out, even if it's not specific to your work. In a theoretical paper on algorithmic fairness, you might caution against overreliance on mathematical metrics for quantifying fairness and examples of ways this can go wrong. 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 different stakeholders that could be impacted by your work. It is possible that research benefits some stakeholders while harming others. Pay special attention to vulnerable or marginalized communities.
- 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 should 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).
- Did you describe the limitations of your work?
 - You are encouraged to create a separate "Limitations" section in your paper.
 - Point out any strong assumptions and how robust your results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). Reflect on how these assumptions might be violated in practice and what the implications would be.
 - Reflect on the scope of your claims, e.g., if you only tested your approach on a few datasets or did a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
 - Reflect on the factors that influence the performance of your 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 able to be reliably used 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.
- 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 strongly are encouraged to discuss the relationship between your results and related results in the literature.
- Did you include the code, data, and instructions needed to reproduce the main experimental results?
 - The instructions should contain the exact command and environment needed to run to reproduce the results.
 - 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.
- 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.

- Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)?
- Did you include the amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)?
 - Ideally, you would provide the compute required for each of the individual experimental runs as well as the total compute.
 - Note that your full research project might have required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper). The total compute used may be harder to characterize, but if you can do that, that would be even better.
 - You are also encouraged to use a CO2 emissions tracker and provide that information. See, for example, the [experiment impact tracker](#) (Henderson et al.), the [ML CO2 impact calculator](#) (Lacoste et al.), and [CodeCarbon](#).
 - If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
- If your work uses existing assets, did you cite the creators? 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.
- Did you mention the license of the assets?
 - 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.
- Did you include any new assets either in the supplemental material or as a URL?
 - During submission time, remember to anonymize your assets. You can either create an anonymized URL or include an anonymized zip file.
 - If you cannot release (e.g., the asset contains proprietary information), state the reason.
- Did you discuss whether and how consent was obtained from people whose data you're using/curating?
 - For example, if you collected data from via crowdsourcing, did your instructions to crowdworkers explain how the data would be used?
 - Even if you used an existing dataset, you should check how data was collected and whether consent was obtained. We acknowledge this might be difficult, so please try your best; the goal is to raise awareness of possible issues that might be ingrained in our community.
- Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content?
 - There are some settings where the existence of this information is not necessarily bad (e.g., swear words occur naturally in text). This question is just to encourage discussion of potentially undesirable properties.
 - Explain how you checked this (e.g., with a script, manually on a sample, etc.).
- 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.

- Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable?
 - Examples of risks include a crowdsourcing experiment which might show offensive content or collect personal identifying information (PII). Ideally, the participants should be warned.
 - 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.
- Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? First, provide the amount paid for each task (including any bonuses), and discuss how you determined the amount of time a task would take. Also include discussion on how the wage was determined and how you determined that this was a fair wage.