A Guide to REAL ML

Guided activities for recognizing, exploring, and articulating limitations in machine learning research.

Introduction

What is this tool?
The Recognizing, Exploring, and Articulating Limitations in Machine Learning research tool (REAL ML) is a set of guided activities to help ML researchers recognize, explore, and articulate the limitations that arise in their research. This document acts as an instructional guide and is meant to be completed with the corresponding REAL ML worksheet that you will fill out as you go through the activities below.

What is a limitation?
The process of conducting research is never without some limitations—*drawbacks in the design or execution of the research that present a possible threat to the validity of your findings and claims*. Research rarely takes place under ideal conditions; instead, researchers must frequently navigate practical constraints and unexpected challenges in seeking to answer their research questions. Researchers also make all sorts of explicit and implicit choices in the design and execution of their studies. All of these may affect the conclusions that researchers are able to draw from their work. In many fields, limitations are understood to be an inherent part of research and researchers are expected and encouraged to disclose and discuss them. Doing so is viewed as critical to maintaining the integrity of a field of research and to helping advance collective knowledge.

How do limitations relate to broader impact statements?
While limitations and broader impacts share some characteristics, the two are not the same. Limitations focus on parts of the research process that might threaten the validity of your conclusions; broader impacts focus on the downstream implications of your research when put to some practical use. For guidance on writing a Broader Impact statement in your paper, we suggest you look at the NeurIPS Broader Impact statement or the NeurIPS Paper Checklist Guidelines.

Who should use this tool?
We have developed this tool for use by ML researchers specifically.

Why should I write about limitations?
- Addressing limitations in your research can improve the scientific rigor of your work by ensuring that you are more precise in describing what the research entailed and what claims your research supports.
- Making the limitations of your research explicit can help readers develop a more accurate understanding of the research project, its findings, and the conclusions that you've drawn from these findings.
- Papers that include critical reflection on limitations can improve readers' justified trust in the research findings; papers that lack such reflection can make readers overly skeptical of research claims.
- Imparting a clearer understanding of the limitations of your research can help to avoid inappropriate applications of your research in practice.
- Pro-actively reflecting on limitations in research papers may help to address possible concerns on the part of reviewers, improving the chances of having a paper evaluated correctly and thus being accepted.
- Describing the limitations of your research—and the sources of these limitations—can foster collective scientific progress by highlighting opportunities for future research; disclosing limitations can also mean that others are more likely to build on your work.

When should I use this tool?
This tool is intended to be used during the paper-writing phase of research. You may also find some of the activities useful earlier in your research, including during the development of your research design.

How do I use this tool?

We recommend that you use this tool for one ML research project/paper at a time. You can complete the activities in this tool on your own, with a partner, or with a group of researchers. When working in a group, you can discuss the answers to questions out loud and have someone write down your thoughts and takeaways. It is recommended that you use the REAL ML Worksheet to help structure your note-taking.

How long should I expect to spend on this?

It will take about 45 minutes for you to get through the first two activities in this tool—learning about "Sources and Types of Limitations" and then "Recognizing" some of them in your own work. If you continue onto the "Exploring" and "Articulating" activities, you should expect to spend at least 1 more hour to complete the full set of activities.

How was this tool created?

This tool was created by researchers at Microsoft Research who interviewed researchers from the ML community about their experience with limitations and with various iterations of this tool. Participants' input and feedback shaped the ultimate design of this tool. More information can be found in the associated publication:

Jessie J. Smith, Saleema Amershi, Solon Barocas, Hanna Wallach, and Jennifer Wortman Vaughan. REAL ML: Recognizing, Exploring, and Articulating Limitations of Machine Learning Research. In *Proceedings of the 5th ACM Conference on Fairness, Accountability, and Transparency* (FAccT), 2022.

Where can I find more information?

The latest version of REAL ML and the accompanying REAL ML Worksheet can be found at https://github.com/jesmith14/REAL-ML. You can contact the team at realml@microsoft.com.

Sources and Types of Limitations

~15 Minute Activity
Familiarize yourself with common sources and types of limitations described below and start to think about ones that you may have encountered in your work.

Limitations can arise in ML research for various reasons including unavoidable constraints, unforeseen challenges, and explicit or implicit decisions. To help get you in the right mindset to think through the limitations that exist in your work, it can be useful to become familiar with these sources of limitations.

In this activity, your task is to note the potential sources of limitations that you may have encountered in your research. Read through the following prompts and answer the questions on your own or with a group. We recommend that you write these down so you can refer to them in later activities.

Unavoidable Constraints
What constraints might have impacted this research that were out of your control? For example:

- Time constraints (e.g., not enough time to run experiments on the entire dataset, had to complete the research before funding was terminated)
- Resource constraints (e.g., lack the desired technical infrastructure to perform the study; not enough money to pay for desired compute power)
- Lack of access (e.g., unable to obtain access to a dataset that would have been more relevant to the research question, not able to secure necessary license to use certain code)

Unforeseen Challenges
What unforeseen challenges did you encounter in this research that might have resulted in limitations? For example:

- Experimental failures (e.g., your dataset was too sparse to create a meaningful model, your approach didn’t perform as well as previously set benchmarks)
- Negative results (e.g., the results of your experimental study do not confirm your hypothesis, a theorem you initially set out to prove turned out to be false)

Implicit and Explicit Decisions
What decisions did you make in your research? Which decisions were more likely to lead to limitations than others?

Decision-making points

The following table gives examples of common decision-making points that you may have encountered during your research:

Decision-making points	Examples
Composition of the research team	Demographic features (e.g., race, gender), disciplinary training (e.g., computer science, medicine), epistemological perspectives (e.g., Bayesian vs. frequentist), or other researcher characteristics that can influence the approach to research and interpretation of the findings
Related work	The specific fields with which your current study is engaging and which may shape your research; prior work to which your current work is responding; prior work upon which your current work is building
Problem formulation	The general problem that motivates the research; the specific research questions developed to get at that problem
Formalism of the problem	Mathematical statement of the problem that your study is trying to address; technical assumptions (e.g., i.i.d. data points)
Technical approach	Learning algorithm; statistical model; hyperparameter choices
Theoretical claims	Theoretical guarantees such as error bounds; analyses of computational complexity; mathematical derivations
Datasets	The collection, curation, and selection of datasets; the use of particular datasets for training or evaluating
Empirical evaluation setup and metrics	Experimental setup including approaches to be compared, metrics, parameter settings; research subjects
Ablation studies	Setup for ablation studies, including components removed and metrics

Types of limitations

These unavoidable constraints, unforeseen challenges, and explicit or implicit decisions can result in a range of limitations, a list of which we provide in the following table:

Types of Limitations	Probes to Uncover Limitation	Examples
Fidelity	How faithfully do the formalism of the problem, the technical approach, and the results map onto the motivating real-world problem that drives the work?	The formalism of the problem includes so many assumptions that results are not clearly applicable to the motivating problem in the real world. There are large gaps between the data/model/metrics and the motivation. The training data was labeled even though similar real-world data is not usually labeled. The distribution of the data is different than what you would encounter in the real world.

Generalizability	To what extent do the results hold in different contexts? How broadly or narrowly should the claims in the paper be interpreted? How broadly can the technical approach be applied across domains?	Model was developed for a particular scenario and does not apply to other scenarios or contexts. Only a single dataset, small dataset, or datasets from a specific domain were used. If taken out of context, the results of this study could be misleading.
Robustness	How sensitive are the results to minor violations of assumptions (e.g., small tweaks to mathematical model, metrics, hyperparameters)?	Two equally reasonable ways of formalizing the problem would lead to dramatically different takeaways. Adding a small amount of noise in the data dramatically reduces accuracy. Selecting different reasonable metrics would lead to different outcomes or tradeoffs for the experiments.
Reproducibility	To what extent could other researchers reproduce the study?	Researchers provide details on parameter settings used but cannot share code or data because they are proprietary. The code and data are shared, but no guidance about running the experiments is given.
Resource Requirements	Is the technical approach computationally efficient? Does it scale? What other resources does the technical approach require?	Approach works well with a couple thousand training examples but cannot handle a couple million. Technical approach requires specialized hardware.
Value Tensions	Are some values (e.g., novelty, simplicity, high accuracy, low false positive rate, ease of implementation, interpretability, efficiency) sacrificed in pursuit of others?	The model has high accuracy on a test dataset but is a black box and hard to interpret. The model is optimized to favor certain kinds of errors over others (e.g., false positives are favored over false negatives). Seeking voluntary and informed consent from human subjects leads to selection bias in the dataset.
Vulnerability to Mistakes and Misuse	How sensitive are the results to human errors, unintended uses, or malicious uses?	System operators are liable to misinterpret results without sufficient training. There is a serious risk if the technique, which was developed in one context, is applied in others for which it is not suited.

Recognize

~30 Minute Activity
Start to recognize limitations in your research and fill in the first 3 columns of your worksheet as you go.

In this activity, you will start to recognize the limitations in your research project. Remember that all research has inherent limitations. Identifying and acknowledging these limitations can be beneficial for your work and for the ML research community. Don't be afraid if your list of limitations is long!

To complete this activity, you will need to use the accompanying REAL ML Worksheet. Your task is to fill out the first three columns of the worksheet. In Column 1, list all the limitations you can think of. Refer to the "Types of Limitations" in Table 2 for help. For each limitation, denote its "type" in Column 2 and its "source" in Column 3, as introduced in the previous section of this tool. Note that one limitation might have multiple sources; this is especially common for limitations that are categorized under the "Generalizability" and "Robustness" types.

Still unable to come up with a limitation?
If you are having difficulty thinking of a limitation, recall all of the different sources that you previously identified. There are always limitations to uncover, whether or not you decide to put them in your final paper.

Explore

~10 Minutes Per Limitation
Use the prompts below to explore different aspects of the limitations you identified that may be important to include in your paper.

In this activity, you will read through a series of questions to help you uncover and explore various aspects of the limitations you identified in the previous activity that may be important to include in your research paper. To complete this activity, you will again use the accompanying REAL ML Worksheet. For each limitation you identified in the previous activity, go through all the questions below and fill in the corresponding columns in the worksheet. That is, you will answer each question (column) one limitation (row) at a time. Take a look at the examples in the table for guidance.

Alternatives
If the source of this limitation was a decision that you or your research team made, what are the pros/cons of this decision? What alternative decisions could you and/or the research team have made instead?

If the source of this limitation was practical constraints or unforeseen challenges, what alternative research designs could you have explored if those could have been avoided? What are the pros/cons of those alternatives?

Write down your biggest takeaways in Column 4 of the worksheet.

What Should People Know?
What are the impacts of this limitation on your study's results and claims? What do different readers of your paper need to know about this limitation and its impacts?

- What might reviewers need to know (e.g., what did you do to minimize the possible impacts of this limitation)? This category is a great opportunity to prepare for reviewer feedback!
- What does the ML community need to know to understand and trust the claims of your work? What do future researchers building on this work need to know?
- What do practitioners who might build on this research need to know?
- What do people who are using or are affected by the systems these practitioners build need to know?

Write down your biggest takeaways in Column 5 of the worksheet.

Repeat this activity for each limitation you identified in the worksheet. It is recommended that you spend at least 10 minutes on this activity for each individual limitation you would like to explore. Once you have completed each row, move on to the next section.

Articulate

~25 Minute Activity
Draft the limitations section of your paper using the worksheet you filled out and considering the tips below.

Now it is time to draft the limitations section of your research paper! For this activity, you will take all the information you have written in your REAL ML Worksheet and coalesce these ideas into a useful narrative. What you write here does not need to be "camera-ready" quality, but it should include details of your limitations that (1) you would feel comfortable including and (2) you think would be useful to include for the audience of your final paper.

Write your draft in a document or on a piece of paper so you can use it in your final research paper. There is no "right" or "wrong" way to do this. If you are feeling stuck, refer to the tips and tricks below.

Tips and Tricks for Framing Your Limitations

Speculating about your claims?
If you encounter a limitation where you aren't sure of the impact on your findings, begin by reporting what you do know. Then, if you choose to speculate about something unknown, be explicit about your speculation. For example: "we hypothesize that our solution could be used in these other ways, but future work should provide verifiable evidence for these."

Running into page limits?
Consider focusing on the limitations, or details of your limitations, that you think may have the biggest impact on the validity of your findings. Additionally, think about your primary audience(s) and what is most important for them to know to accurately interpret your research claims. In many research venues, you can include further details or even your completed REAL ML Worksheet in an appendix to your research paper.

Worried that a limitation is grounds for rejection?
We understand that authors might fear that complete honesty about limitations could lead to their papers being rejected. As noted in the NeurIPS paper checklist, you should keep in mind that your paper might be more likely to be rejected if reviewers discover limitations in your paper that you didn't acknowledge ahead of time and explain why your research still has merit despite these limitations. Recognize that transparency around limitations also helps to preserve the integrity of the ML research community, foster collective scientific progress, and ensure appropriate applications of research findings in practice.

Want to discuss a limitation's broader impact?
You might have noticed that certain limitations have the potential to lead to serious downstream harms, especially if your research is misapplied or applied maliciously. We recommend that you separate this discussion of broader impacts from your limitations section. For guidance on writing a Broader Impact statement in your paper, we suggest you look at the NeurIPS Broader Impact statement or the NeurIPS Paper Checklist Guidelines.