



# Journal of Visualization and Interaction

JoVI review snapshot created at Dec 10, 2025 13:05.

**Issue #5 (open): [REVIEW] PAC Learning Or: Why We Should (and Shouldn't) Trust Machine Learning**

@Chenglong-MS on Apr 12, 2024 16:34: [opened]

@Chenglong-MS on Apr 12, 2024 16:34:

## Conflicts of interest

- ☒ I declare that I have no known conflicts of interest with the authors.

## Reviewed version

fef096a

## Review

### Review

This is a fantastic paper demonstrating how interactive plots can help introducing PAC learning to the general audience.

First, I'm in full support of this paper. I believe the interactive paper like this will be very beneficial in teaching and education towards general audience. Let me directly jump into some suggestions I think might help with paper presentation.

- The first paragraph of intro does not give a sense of "game" for the audience: there is no clear explanation of the user task, the interactive plot, and the user's goal. I would suggest (1) clarifying the task and goal e.g., "you will try to find the best rectangle that defines the concept 'men of median height and weight' with certain number of examples, and your correctness is measured by the metrics FP/FN/TP/TN", (2) explain the interaction that the user needs to draw rectangles, (3) explain the difference between training and testing samples (in particular, I highly suggest labeling training and testing samples differently, so that the reader can clearly see what points they use to derive their bounds, and what are new points used in tests).
- The "Why is it important" section is a little too generic. While I understand it's necessary to relate existing AI advancement like LLM so that general public would better appreciate the article, it's better to start with some smaller yet more related real world problem (in relation to PAC learning). E.g., regression/classification tasks that are used in business, banking, real-estate.
- In the PAC explanation section, it would be a good idea to label variables in the plot directly to help reader leverage the plot to understand the derivation process.
- The title "Assuming the worst" is somehow misleading, since it's more or less related to adversarial examples. In fact, it's very common in practice to have these issues. Furthermore, also a good idea to provide more examples here about why these non-ideal situations would happen. For example, you can use medical example to explain why samples are non IID, since patients won't get certain tests unless they are in severe situations.
- There could be an additional section before conclusion to point out what are further reading beyond PAC learning, and how it is related to existing ML practices.

## Additional suggestions

Here are some suggestions that are not required (since they may require additional programming), but they could be nice to have.

- I kinda hope the sense of the "game" can be more highlighted in the interactive chart. For example, in the model class selection, it could be fun to let readers choose what model class they like.
- Before conclusion section, it could be nice to have several interactive examples to let the user play with some real data, to understand the gap between the theory and practice, but also highlight why theory would be helpful. If I were you, I might build a few examples that are combinations of different conditions (e.g., non-iid, small data // iid, small data // iid, non-real model class) without telling the author, and ask them to play for prediction.

## Technical issues:

- When hitting "TEST!" with a rectangle, an error message `Cannot read properties of null (reading 'x')` would show up and prevents further interaction. Would suggest showing an error message you need to draw a rectangle that shows your concept guess before hitting TEST!.
- The layouts of the paper is some what wierd, since the left texts and righthand side plots are not always aligned. The uses of blank space between small sections are nice for the screen but they make navigation difficult (related bullet points are too far apart). I think it would be a good idea to include some visual cue which bullet point / section the right hand side plots links to.

## Conclusion

I'm fully in support of the paper; some revisions would be helpful to improve the paper quality.

## Openness/Transparency

not applied since the paper is towards interactive example.

## Submission categories

- ☒ Registered Report
- ☐ Replication Study
- ☐ Empirical Research - Quantitative
- ☐ Empirical Research - Qualitative
- ☐ Systems or design research
- ☐ Commentary
- ☐ Systematic Literature Review

## Suggested outcome

Minor revisions: this paper requires some smaller changes, after which I am confident I would be able to endorse it.

## Requested changes

Please refer to the reviews.

Requested changes are under review tab. Optional reviews are under additional suggestions tab.

ORCID

No response

@dylancashman on Apr 22, 2024 15:48: Hi @Chenglong-MS, thank you for your review! I appreciate all of your comments and will work on them soon.

@dylancashman on Dec 01, 2024 02:25: Improved introduction in <https://github.com/journalovi/2024-Cashman-PAC-learning-game/pull/11>

@dylancashman on Dec 01, 2024 08:25: Cleaned up bug with Error when no rectangle was drawn in [#12](#)

@dylancashman on Dec 04, 2024 02:47: Thank you for your review! I wanted to provide a little information on what I addressed in my revision.

- 1. The first paragraph of intro...

I added some more context in the introduction to explain and motivate the game.

- 2. The "Why is it important" section is a little too generic.

I added a more simple example of spam filtering, since it is an ML example that everyone has experienced.

- 3. Labeling variables in PAC Learning example

There were already labels for T and T', but they were very hard to see, so I updated the font size to emphasize them more.

- 4. Assuming the Worst

I did update this section with a few additional examples, including one about a medical use case with drift.

- 5. Additional reading

I did not add any additional reading section because it was not listed as a required revision in the meta-review. However, I do think that the listed references provide some good pointers : people wanting to read more.

I appreciate the additional suggestions for making the game more interactive. However, as these suggestions were not listed as required revisions in the meta review, I did not incorporate them in this revision.

In addition, I did address the TEST! bug that you identified. The white space is somewhat limited because of the format of the article published by idyll.

@Chenglong-MS on Jan 17, 2025 21:05: Thanks a lot for the revision! I checked revisions with respect to my prior comments, and they are completed very well! I especially enjoy the "Why is it important" revision to make it mo clear.

Here are some comments from the UI perspective (both optional):

A little technical issue is that while I scroll down the document, this "datum not found" will appear. It doesn't seem to get in the way of reading the document (it can be refreshed after clicking "try it now"), but might be something worth to fix to avoid confusion.

A LITTLE LUNCH, AS A TREAT

However, suppose we are given some information. Here, we are given 10 points, with 5 inside the target rectangle and 5 outside. If you guess a rectangle, can you make any guarantees about how close you are to the ground truth?

Consider the various strategies that you might use to minimize your error. Do you tightly wrap the green examples, or do you leave some space around them to allow for data you haven't seen yet? Which strategy generally works better? Which strategy works better in the unlucky case, where the sampled data doesn't provide much information due to bad luck?

Up to this point, you have been playing the machine learning algorithm. You have all the magnificent faculties of the human mind, letting you change and adapt your strategy as you encounter new examples. But machine learning algorithms tend to be much simpler, and typically more stuck-in-their ways, because they have to be well-defined.

datum is not defined

↓

Also, I think it could be a good idea to add a little animation for the interactive panel when it switch sections to inform the user that the right side panel is refreshed/updated to the new game. (The title did update, but a little more visual cue would be super). Something like this could be handy:

```
```refresh-fade { animation: refresh 0.5s ease-in; }

@keyframes refresh { 0% { opacity: 0.5; } 100% { opacity: 1; } } ```
```

@Chenglong-MS on Jan 21, 2025 18:08: Decision: Endorse: I am willing to endorse this paper, with at most minor copyediting

@dylancashman on Sep 30, 2025 20:46: Thanks for the updated comment! I added a transition for the text at the top of the game. I was not able to reproduce the "datum is undefined" error, but there are some null checks in the code as is, so maybe that was from a previous build.