Journal of Visualization and Interaction

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

Issue #7 (open): [REVIEW] PAC Learning Or: Why We Should (and Shouldn't) Trust Machine Learning

@MarkoAngelini on May 20, 2024 08:42:

[opened]

@MarkoAngelini on May 20, 2024 08:42:

REVIEW

This paper presents an interactive game representing the types of tasks solved by machine learning algorithms. It supports the definition of Probably Approximately Correct (PAC) learning, illustrating proof of PAC learnability for Empirical Risk Minimization (ERM). Then, it shows three vulnerabilities for ERM.

The paper briefly introduces the problems of machine learning and the Four Germans game, a game where is essential to estimate the rectangle that effectively splits correctly ALL the points of a binary classification, being guided only by a subset of the points (eventually even zero for a complete guess)

Through a set of incrementally complex versions of the game (starting from zero support from training data to full support) it allows for review of some of the basic concepts of Machine Learning.

The first version of the game presented is without any point: just a guess, to discuss the extreme case.

Next, a version with ten fixed points is given.

This is helpful for introducing different heuristics for estimating the containing box /its general version, not the one specific for the visible points): Tightest fit (the tightest possible rectangle is chosen), Loosest Fit (the algorithm chooses the loosest possible rectangle), and Maximum Margin (a midpoint between the two, which tries to maximize distances from both)

These incremental steps help the reader get in touch with the problem and, with partial interactivity supported, create their own mental model through active testing.

This part is interleaved with theoretical concepts and definitions, in particular for the PAC learning algorithm and PAC-learnable spaces and its implication on the correct choice of an ML algorithm for a specific problem, based on a bound error Ïµ and a probability of the error being bounded 1âˆ’Î´.

Overall, the authors argue that by proving that an algorithm is a PAC learning one it is possible to avoid applying a wrong ML algorithm for a specific task and get a false sense of accuracy by the â€œby-defaultâ€ applied Empirical Risk minimization on the training data. On the other hand, if this is not the case, the analyst can receive a false sense of confidence in learning results while effectively not being in a correct generalization case. To explain this concept, the authors first consider that the tightest fit is always contained in the target rectangle, and then partition this part into four strips.

Then, first focusing on the top strip, it is shown how through geometric properties, error is bound by exactly (1- Ïµ /4)^M. Multiplying it by four (the four strips) and solving for Ïµ the original implication is demonstrated (the error is bounded for unseen examples).

The process then focuses on how many samples are needed to find the correct estimated box. This equals to solving the previous equation for M, with the result that M is not infinity and so the problem is PAC-learnable as the definition dictates.

The environment then moves on to describing several assumptions taken, not all the time valid for real applicative cases, without which the presented demonstration and properties are no more valid. The first is the violation of independent and identically distributed data for sampled data, without which the prototype shows things are not working correctly anymore.

Similar considerations are valid for the very well-known case in which the distribution of training data has nothing to do (or in general is not the same) as the one for the test dataset. The interactive example, taken in the first extreme case, shows that no guarantee can be given to the estimation of the box.

A third assumption is relaxed in the form of the â€œcontaining boxâ€ to guess (a generic polygon or even more generic). The test shows the problem in practice with an elliptic box.

This part ends with providing more advanced references on PAC learnability and some limitations in its applicability.

Overall, in all these cases, the interactive examples help at grasping the real problem, even if sometimes just at the surface level. No matter what, this can be good for teaching or getting introduced to the matter one step at a time with more engagement.

The last part, which opens more broadly to visual explanations for AI is the weakest as it just presents static text. It could have been interesting, in line with what has been presented up to that moment, to provide an interactive glimpse of what it means (i.e., an example taken from the referenced Ph.D. thesis).

The text flows well, with some minor corrections reported at the end of this review.

Revisions/Questions (Q):

Section 1: improve the points explorer/testing environment Q1: When I played the game, the first thing I did was wait a little for some green points to appear, identify a potential box, and then draw it to the border of limiting red points. Then I pushed the TEST button: what I saw was that at that moment many of the green points I used for reference were not displayed anymore, while new green points appeared (with the green box not perfectly overlapping, but that is correct as intended). Why are the already visible points removed? This creates some confusion in the user. Those points should be kept in the view while the new ones could be added, simulating the acceleration of the remaining part of the dataset passing through the model.

Section 2: clarify the assumptions for the four side of the box Q2: In estimating the toughest fit error, why is it implied that each strip size must be lower than epsilon/4? Does it not imply that the contribution to the error is equal for each strip? Would it be more correct to say that the sum of the errors must be lower or equal to epsilon more correct, allowing different contributions to the error by different strips?

Or more generally speaking, does this assumption affect the final equation for estimating M, being a parameter for it, or instead is it the M estimation agnostic to it apart a constant term?

Section 3: clarify the hosting of the final prototype (probably JoVI as it is) Q3: the system right now is a packaged web application to launch locally in order to make it work. Do the authors plan to host it somewhere? The reviewer asks it, as, for example, launching it in a completely local environment without an internet connection breaks the formula formatting and some parts of the layout (tested with both Chrome and Firefox browsers) while the backend and interactive examples continue working flawlessly. Would it be good to provide even a link to the hosted version (if any) or details about its hosting?

Section 4: improve the explainability coordinating the points classes with the points selections Q4: this is more a suggestion (not mandatory to comply with): would it be good to make also the four classes of outcomes (TP, FP, TN, FN) interactive, selecting at any stage the subset of points corresponding to them? It could help the reader quantifying the introduced errors also in the â€œgeometric spaceâ€

MINOR revisions:

A) The text size of the paper is quite small even on big screens and difficult to read. Both the manuscript text and the interactive environment on the right could better fit the screen size vertically to make the elements more visible (The reviewer checked that the environment is responsive, so the problem seems more correctly exploiting the dynamic screen size readings). This is true even for the PAC-Learning section.

B) The â€œLoosest Fitâ€ example seems too conservative on the left and right sides (there is still some area that can be covered with respect to the most right and most left red points).

C) sometimes the â€œinteractive partsâ€ are covered by textual parts, in particular by definitions. It seems a problem of overflow management and visibility.

D) The transitions of the interactive part through the different stages by scrolling down and up functions in a very hectic one, sometimes freezing in a wrong state with respect to the part of the text chosen. While the blue buttons restore the correct situation, the risk could be for a not-savvy reader to misinterpret that part.

E) the â€œbig arrowsâ€ used to guide the reader in scrolling down show a bad presentation (leaving a lot of vertical space and not working as automatic scrolling to the next anchor, at least in the reviewer experience). The reviewer suggests making them better distanced, more visually appealing, and interactive to automatically reach the next anchor.

Decision:

Overall, the reviewer believes that the work is valuable and well-constructed. On the other hand, some revisions are needed in terms of the written part, its interactive aspects, and in their fluid combination, to make it of the quality of the JoVI Experimental track. For these reasons, the reviewer argues for a major revision score and asks the authors to consider the requested revisions, try to implement them, and/or contact me for any clarification or discussion on their meanings.

MINOR comments/typos::

1) the underlying phenomonen â€“ > the underlying phenomenon

2) (here error defined as the total area that is only inside one box, i.e. false positives (in our proposed concept, not in the ground truth concept), or false negatives (in ground truth concept, not in our proposed concept)) â€“ > please avoid nested parentheses

3) probably approximately correct (PAC) learning,-- > please report the primary source instead of Wikipedia page (it is listed already in Wikipedia as the original 1984 paper)

4) precictions â€“ > predictions

5) demonstrate why fully-automated approaches will never be able to be error-free if these assumptions are validated. â€“ > Is this sentence a little too strong?

Why this is important

6) effected by them â€“ > affected by them

7) called called Genderâ€¦ â€“ > called Gender

8) Risk Minimization â€“ > This link is broken (page not found)

PAC learning

9) Blumer et. al. â€“ > Please add a hyperlink to the reference

10) labeled trainign â€“ > labeled training

11) Sometimes formulas are not correctly rendered, e.g., â€œâ€¦ than ðœ–Ïµ, with probability at least 1â^'ð›¿1â^'ð›¿â€¦â€

12) with much more complex data and intricate, high-dimensional processes -- > this sentence seems a little too vague and â€œflashyâ€. What are â€œhigh-dimensional processesâ€?

13) topiccan -- > topic can

14) â€œthis article on its Github repository.â€ â€“ > the link does not bring the user to the Github repo but to this paper: https://www.liebertpub.com/doi/full/10.1089/big.2016.0051

@dylancashman on Dec 01, 2024 08:26:

Clarified which points are drawn in #12

@dylancashman on Dec 04, 2024 02:28:

Thank you for your review! Adding some clarifications about the revisions that were made.

> Section 1: improve the points explorer/testing environment

I added a counter for the number of training points and the number of testing points shown. The training points disappear because they are not part of how the model is evaluated - it's to illustrate what happens when we put a model out into the wilderness.

> Section 2: clarify the assumptions for the four side of the box

The proof that is illustrated is taken from the cited literature, so I am somewhat constrained in what I can change, but I did change the description of the proof somewhat to point out that the errors are based on uniform sampling of the space, and so the area of the strips correspond to probability of errors.

> Section 3: clarify the hosting of the final prototype (probably JoVI as it is)

For the internet connection issue, I am happy to provide a static build that is hosted on JOVI. I'm not sure how that happens though. The project is fully buildable and does not rely on any backend server.

> Section 4: improve the explainability coordinating the points classes with the points selections

While I like the suggestion, I ultimately didn't change the interactivity of the game because it was not mentioned as a necessary revision in the meta review.

Beyond the above revisions, I made some grammatical and typo-related fixes. I also spaced out the sections a little bit to make the transitions less jarring in the proof stage.

@MarkoAngelini on Mar 10, 2025 11:01:

Thank you for working on the suggested revisions.

Looking at them in more detail:

Section 1: the introduced counters help in keeping track of the relation. Additionally, nut maybe this is a connection dependant, the original training points tend to stay for a fraction of time (a couple of seconds) still visible after clicking the â€œtest!â€ button, not creating anymore the original confusion about them abruptly disappear. I consider this revision addressed.

Section 2: Ok, I understand you followed the original formulation. The added textual explanation helps clarify the subject. I consider the revision addressed

Section 3: Fine that now a static build is hosted on JoVI. It solved the problem of looking for external connections when an internet connection is not available. I consider the revision addressed

Section 4: I accept that, due to its not mandatory nature, the author chooses not to implement the revision. At the same time, I still believe it may help the interactivity and detailed analysis of differences between the training and test data on the classic four classes of outcomes which otherwise, as of the current version, are just reported as summaries without any inspection capabilities. At the same time, given the optional nature, I consider the rebuttal acceptable and the revision rebutted

All the reported minor revisions and points have been successfully addressed.

On the contrary, some quite minor problems remain somehow â€œdisturbingâ€ mostly the usability of the submitted work and not anymore its correctness.

Minor revision 1: I confirm that as of today, clicking directly on a lot of functionality for testing (e.g., loosest fit, a little lunch as a teat, etc.) returns the â€œdatum is not definedâ€. As much as I believe this is just dependent on where the analysis was, it should be made more clear what actions the user needs to perform to exit the error condition (i.e., making the error message informative on how to solve it, like â€œPlease select an area firstâ€) and more generally minimize the condition it happens (maybe keeping the previous selection made by the user, and informing it about that?). This error seems to prevent the execution of several parts of the environment and so check the eventual modifications made by the author.

Minor revision 2: I noticed that in any status of the interactive environment, after clicking on the â€œtest!â€ button, the user can accidentally still select an area, destroying the original â€œgreyâ€ area representing their original â€œtraining data areaâ€ specification. Would it be possible to draw that original selection with a dashed rectangle to keep it persistent, avoiding reusing the active area of selection which can be simply destroyed by accidents clicking again the cross-hair on the screen? This in my opinion is very important for the sections about training-test mismatch, focused exactly on showing the amount of the mismatch and potentially where it is localized

Minor revision 3: sometimes in the local version the axes disappear from the screen (I am using Google Chrome version 133.0.6943.127). It happens mostly during scrolling. Even after hitting the â€œrefresh buttonâ€ the problem seems to persist. I tried also Firefox and the problem is also present for it (to test it is sufficient to click on the first button of the â€œFour Germans Game: Find my rectangleâ€ section.

Minor revision 4: I second what I read about another reviewer asking for â€œanimationâ€ to convey changes between different configurations of the interactive environments. I also suggest using one to make the interactive environments disappear and reappear during the scrolling under the definition of PAC learnability. The environment scrolling under this part is not very effective and, given the definition completely covers it, is unnecessary.

Very minor revision 1: maybe it is my problem, but it seems anchors do not work anymore nor in the local version or hosted version. Please check

Decision: important aspects of previous major comments have been addressed. Some minor points remain that can be quite quickly addressed to complete the work. I suggest a minor revision decision asking the author to address the requested minor points and after that, the work can be accepted.

@dylancashman on Jul 11, 2025 04:10:

Thank you for your careful review and feedback. I have made the following changes:

> Minor revision 1: I confirm that as of today, clicking directly on a lot of functionality for testing (e.g., loosest fit, a little lunch as a teat, etc.) returns the â€œdatum is not definedâ€. As much as I believe this is just dependent on where the analysis was, it should be made more clear what actions the user needs to perform to exit the error condition (i.e., making the error message informative on how to solve it, like â€œPlease select an area firstâ€) and more generally minimize the condition it happens (maybe keeping the previous selection made by the user, and informing it about that?). This error seems to prevent the execution of several parts of the environment and so check the eventual modifications made by the author.

- This error has been fixed. It wasn't able to calculate some things without anything selected, but I have fixed it to function with nothing selected.

> Minor revision 2: I noticed that in any status of the interactive environment, after clicking on the â€œtest!â€ button, the user can accidentally still select an area, destroying the original â€œgreyâ€ area representing their original â€œtraining data areaâ€ specification. Would it be possible to draw that original selection with a dashed rectangle to keep it persistent, avoiding reusing the active area of selection which can be simply destroyed by accidents clicking again the cross-hair on the screen? This in my opinion is very important for the sections about training-test mismatch, focused exactly on showing the amount of the mismatch and potentially where it is localized

- This was actually intentional - this way, after a user has seen the answers, they can redraw their candidate to see that they are unable to completely remove error in some cases.

> Minor revision 3: sometimes in the local version the axes disappear from the screen (I am using Google Chrome version 133.0.6943.127). It happens mostly during scrolling. Even after hitting the â€œrefresh buttonâ€ the problem seems to persist. I tried also Firefox and the problem is also present for it (to test it is sufficient to click on the first button of the â€œFour Germans Game: Find my rectangleâ€ section.

- This was a bug, and it is now fixed.

> Minor revision 4: I second what I read about another reviewer asking for â€œanimationâ€ to convey changes between different configurations of the interactive environments. I also suggest using one to make the interactive environments disappear and reappear during the scrolling under the definition of PAC learnability. The environment scrolling under this part is not very effective and, given the definition completely covers it, is unnecessary.

- I added a cleaner transition of the board fading out as you get to the definitions

> Very minor revision 1: maybe it is my problem, but it seems anchors do not work anymore

nor in the local version or hosted version. Please check

- I have confirmed that they are functional on Google Chrome. I believe this was a side effect of other errors previously mentioned.

@dylancashman on Jul 11, 2025 04:10:

[closed]

@MarkoAngelini on Sep 20, 2025 10:31:

I thank the author(s) for the commendable work done in revising their work following the reported suggestions. With respect to my latest minor revision review, I consider Minor revision s 1, 3 and 4 correctly addressed. I understood also the rationale behind the behavior evidenced in Minor Revision 2, whcih is reasonable.

In light of these considerations, I do not have any additional requests and I argue for accepting the work in its current form