Journal of Visualization and Interaction
JoVI review snapshot created at Sep 29, 2025 13:29.

Issue #10 (open): [REVIEW] R2 Review

@joviewer-xyz on Feb 28, 2025 19:53: [opened]

@joviewer-xyz on Feb 28, 2025 19:53:

Conflicts of interest

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

Reviewed version

910f164

Review

In this essay the authors explore a collection of probabilistic graphical models (HMMs, MEMMS, and CRFs) through the specific example of named entity recognition. Throughout, an emphasis is placed on providing graphical explanations for the models considered. While providing visual explanations for complex to understand machine learning models such as these is valuable, I find that this essay lacks a clear contribution to visualization or HCI research beyond existing literature. It is certainly true that providing expositions on topics can be helpful, for instance "Why Momentum Really Works" is an inspirational (and personally influential) example of the genre of online explainer that non-trivially adds useful and interesting visual representations of a complex ML topic. However, this article does not rise to that standard, instead providing a collection of reasonably common UIs which are not synthesized into a larger or more useful lesson.

Strengths:

- Interesting domain: PGMs are complicated and difficult to understand, providing a collection of visualizations that convey a richer intuition for them could be valuable. More over named entity recognition seems like a great toy problem for this as it is realistic and a reasonably common task.
- Apparent coherence of mathematical explanations: While my knowledge of these models is relatively shallow, the mathematical presentation appears coherent. However, I emphasize that my expertise is primarily in visualization and so I am not able to review the machine learning aspects from an expert's perspective, but rather from someone with only a modest amount of training in the topic.

Major Weaknesses:

- **No contribution**: Most crucially missing from this work is a completed contribution that is specific to InfoVis and/or HCI. This review of reasonably standard PGMs could be useful if it offered a richer synthesis or provided more useful exploratory tools that support an enriched understanding of those models. However the visualizations presented are reasonably standard (for instance the various chain/dags throughout are a common teaching visualization for them). To this end, trying to view this from a visualization contribution there are a handful of graphics that might be seen as attempted contributions

1. the DAG chains: these are reasonably standard teaching visualizations for these models (for instance in minor variation from the cited textbook "Probabilistic Graphical Models")
2. the tabular precision/recall heatmaps: while specific to the models, using heatmaps for tabular information such as this is common and does not provide insight specific to this context. I'd add that it was specifically difficult to have to tab between precision and recall on some figures as it made it hard to compare their states.
3. the model inference display (e.g. "name tag predictions by HMM"): while these widgets are

fun to play with they only show model inferences. A richer display might have also highlighted the relative effect (other than just the labels) that each word had on the used model so as to support an intuition about the processes behind them.
4. the sampled population visualization
5. the "Calculating p(A, B, C, D, E, F)" visualization

The last two of these seem most closely centered on visualization as they attempt to support a visual explanation of the studied phenomena. Both of them are well engineered and work in an interesting manner. I specifically liked the tool tip and ability to change generating distribution on (4), however I did not like if any parameter is changed at all then the resample button has to be pressed, which makes it very difficult to use this as an exploratory interactive. Moreover, however, the extremely abstracted nature of both (4) and (5) limits their value as explanations to all but those who are already familiar with the area. I am specifically surprised not to see a richer visualization of HMMs as the hidden layers are so pliable to graphical representation (such as in the wikipedia article about them).

In total, while this essay provides an interesting explainer for this class of models, in its current form this feels like a very good class project, rather than a full research contribution. In order to rise to a contribution I would require a richer collection of visualization/visual explanations that expand beyond simple interactives. Or, more holistically, I'd like to see some type of summative visualization that allows for informative comparison of all of them model types.

- **Unclear audience**: I found myself struggling to figure out who this article was for. The level of assumed knowledge was sometimes high enough (or the explanation given limited enough) that one would assume that the reader already has a non-trivial knowledge of these specific models. Moreover, the work leans on a lot of assumed knowledge of conditional probability and machine learning, which may not be reasonable to assume. For instance, an opportunity missed would be to explain visually the difference between the accuracy achieved and the meaning of the precision and recall. Similarly, the discussion of B/I/O prefix/tags flies by and misses another opportunity for a visual explanation. This is a tough, complex and interesting domain that the authors have chosen and it would be great to have help understanding it.

Minor:

- On the first interactive "Which of these words refer to a named entity?" the dotted underlines make it look like those elements should be interactable, but are not.
- Where did the categories for the labels come from? (eg org/per/loc/misc) Are these standard?

Outcome

I believe that the most appropriate outcome for this work would be a rejection. If i'm not mistaken this is a former VISxAI explainer, and while it does an okay job at explaining its topic it does not do so in a way which yields a contribution in any of the list of submission categories below. There is certainly some interesting thinking present, but, I believe, it is not systematic or repeatable in a way which advances knowledge in HCI or visualization. That said, I welcome the other reviewers who might have more expertise in the area of machine learning to over-ride me.

Openness/Transparency

The code is fully available for investigation and works reasonably well. There are no external research instruments.

Submission categories

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

Suggested outcome

Reject: this paper cannot be fixed to the point where I would endorse it.

Requested changes

NA

ORCID

No response

@joviewer-xyz on Feb 28, 2025 19:55:	[referenced from #[PRE-REVIEW] Review Request R2]
@jsndyks on May 25, 2025 10:35:	[referenced from #[DECISION] Minor Revision - May need more than 1 month and is somewhat dependent on other VISxAI papers & decisions]
@jsndyks on Sep 05, 2025 14:46:	[referenced from #[DECISION] ACCEPT / ENDORSE]
