Why should we teach machines to read charts made for humans?

Jorge Piazentin Ono
New York University
New York City, United States
jorgehpo@nyu.edu

Ray (Sungsoo) Hong
New York University
New York City, United States
rayhong@nyu.edu

Claudio T. Silva
New York University
New York City, United States
csilva@nyu.edu

Juliana Freire
New York University
New York City, United States
juliana.freire@nyu.edu

ABSTRACT

Recent advances in Machine Learning (ML) have enabled the creation of models that can read, parse, and to a certain extent, reason about data visualization charts. These models cannot reach human performance yet, but given the fast progress of the field, they might reach superhuman performance in the near future. We, as HCI researchers who use visualization as a primary way to support human-data interaction, think it is valuable to pose the question: "Does it make sense to make machines read charts made for humans?". In this paper, we portray possible cases where enabling ML models to read visualizations could be beneficial to the community. In particular, we describe three use cases that show how such a model can potentially speed up the visualization design process, and even allow us to come up with novel visualization designs. We also describe possible research directions that can engender fruitful outcomes in the future.

CHI'19, April 2019, Glasgow, UK

KEYWORDS
Chart Data Retrieval, Evaluation Methods, Human-Centered Machine Learning

INTRODUCTION

In recent years, there has been a growing interest in combining Machine Learning (ML) and Human-Computer Interaction (HCI) with the goal of understanding each of these two fields in more detail. As a result, two major research trends have appeared. On the HCI side, practitioners have investigated how Visual Analytics systems can be used to interactively build, explain and compare models [1, 6, 9]. Conversely, ML researchers have tackled the problem of building automated systems that are able to extract information from visualizations and infographics, partially emulating the behavior of a human [2, 5, 7]. While the former was well received by the community, leading to more democratic and interpretable ML tools that help address the increasing concern regarding bias and fairness [4, 19], the latter remains a controversial topic.

Using ML to reason about graphical data representation is a challenging problem. For starters, there is a big difference between the graphical perception capabilities of humans and state of the art ML algorithms. ML models are able to perform significantly better than humans on simple perceptual tasks, such as retrieving position, length and angle from an image. However, not even state of the art models can solve complex tasks involving relations between the data, such as identifying minimum and maximum values of bar charts [5, 8]. The second challenge is making sure the automated model accurately emulates the performance and reasoning of the human. Even if both human and the model performed equally well, we have no way of ensuring the model is replicating the way a human would read the visualization, as that would require a much deeper understanding of perception and of the brain than we currently have. Finally, one can imagine that convincing an HCI researcher to use such systems on their data visualizations would require a lot of effort, as humans have always been at the core of visualization research [13].

As challenging as it may be, designing ML systems tailored to read visualizations could lead to many advantages to the HCI community: 1) these systems would make research more reproducible, by enabling readers to retrieve the data from charts and publications; 2) they would simplify the evaluation of visualizations, making the process faster and cheaper; and 3) they could be used to automatically generate charts from data, based on the visual encodings it has learned. In this paper, we discuss the possible applications of ML models that automatically read and parse data visualizations. In other words, *"Why should we teach machines to read charts made for humans?"*. We argue that while there is much to be gained from using ML to better understand data visualization and HCI, nothing can replace a real life human.

BACKGROUND

The use of Machine Learning to retrieve information from data visualizations is a challenging task, but substantial progress has been made in the past few years. We have identified works in three major areas: perception studies [5, 16], Visual Question Answering (VQA) [7, 8], and chart data retrieval [2, 14].

Perception Studies: Haehn et al. [5] studied how Convolutional Neural Networks (CNN) performed when applied to five graphical perception tasks, including Cleveland and McGill's elementary perception tasks and position-length experiments [3]. They found that CNNs are able to meet or outperform human task performance under limited circumstances, but they are not currently a good model for human graphical perception, particularly for comparison and relation tasks.

Data Visual Question Answering: Visual Question Answering (VQA) is a machine learning area of research that deals with answering open-ended questions about images. Kahou et al. [8] and Kafle et al. [7] extended VQA to answer questions about data charts, and provided baseline results on how well state-of-the-art VQA methods can solve this problem. Their results indicate that while VQA methods can decode the chart structure (e.g., "How many bars are there?"), more powerful models must be developed to reach human-level performance for reasoning questions, such as "Is X the minimum value?". These results are aligned with the perception studies of [5, 16].

Chart Data Retrieval: Charts and graphs are pervasive in newspapers, scientific articles, textbooks and web pages. However, the data underlying the chart is not always available, which leads to problems with respect to machine readability and reproducibility. In order to address this issue, automatic and semi-automatic methods to extract data from bitmap and vector charts have been proposed. For a comprehensive literature review on the topic, the readers are referred to [14]. Related to chart data retrieval is the task of Learning Visual Importance, i.e., identifying the most important graphical elements in charts. These methods can help us better understand how humans read visualizations and infographics, and furthermore, enable us to summarize and restructure charts to facilitate reading. A summary of the state of the art methods can be found in [2].

DISCUSSION

Currently, ML models are not able to read data visualization charts nearly as well as humans do. Our visual system has powerful pattern detection properties that cannot be beaten even by the most advanced ML algorithms, especially when the problem we are trying to solve is not well posed [11]. However, the ML field is advancing very quickly, as shown by methods that can achieve superhuman performance in multiple tasks, such as image recognition [18] and game playing [17]. The authors expect that in the near future, we will have ML models, hereby called *auto-users*, that can read and understand data shown in graphs. As a result, we, as HCI researchers, can start to think about how we can take advantage of these novel methods in order to better understand, design and evaluate our visualizations and graphical representations.

Semi-automatic evaluation of visualizations

Let us assume that there exists a Data VQA model with performance equivalent to a human. If such a model exists, one can also imagine that it could return a metric of *Ease of Use*, how hard it was to extract the information from the chart, and *Confidence*, how confident the model is about the retrieved data (note that most of the current classification algorithms are able to return a confidence score [12], so this assumption is not too far fetched). These metrics can potentially correlate with the quality of the visual encoding in the visualizations, and therefore, could hint at how good the visualizations are in conveying knowledge to a ML model.

This ML-based evaluation would be useful in a pre-user study stage, where the HCI practitioner is iteratively changing and testing his visualizations. In this stage, having the *auto-users* test the system could help in guiding the visualization designer towards the right direction. Furthermore, since these automated testings do not require humans, they would result in cheaper and faster design iterations. However, this process should be done with care, given that the *auto-user*'s and the human's perception might work differently. In order to identify these cases and interpret if this measure of goodness is valid, the vis designer should probe the ML model and investigate HOW it is making its decisions, for example, by using ML model explanations [2, 15].

Ultimately, visualizations have to convey information to people. Therefore, once the HCI designer is satisfied with his charts, a proper user study with humans should be conducted. For this reason, we call this application "Semi-automatic evaluation of visualizations".

Semi-automatic evaluation of interaction mechanisms

Visualization systems usually contain interactions and animations that let the users explore the data of interest with a lot of freedom [11]. Mechanisms such as cross-filtering, pan and zoom are commonly found in visual analytics applications. Could machine learning be used to help HCI practitioners evaluate their visualization interactions? We think so.

With recent advances in reinforcement learning research, we have seen ML models that can learn how to play Go, Chess, and Atari games *tabula rasa*, without any training data or prior knowledge of the game [17]. Similarly to game playing ML models, we believe that it is possible to train a machine to automatically use a visualization system in order to answer specific questions. Given such a model, we could use a workflow similar to the previous section to speed up the evaluation of interactive visualization systems: the HCI practitioner can use the ML model to perform preliminary evaluations of the system, and once they are satisfied with their design, conduct a real life user study of their tool.

Automatic generation of data visualizations

Once we are able to use ML to evaluate visualizations, we can think of even more exciting applications for the *auto-users*: the automatic generation of data visualizations. Given that *auto-user* can compute a visualization score of "quality" for a particular dataset, task, and visual encoding, we can define an optimization problem to find the visual encoding that produces the best visualization automatically. Moritz et al. [10] have used this approach to automatically generate visualizations for data. However, they hard coded theoretical design knowledge to generate rules that guide the visualization building process. We believe that by using an automated evaluation metric to guide the optimization, this process can be improved, and we can discover new and exciting ways to represent data, just like Google's Alpha Zero was able to discover never-before-seen strategies to play Go [17].

CONCLUSION

In this position paper, we have argued that it is important to teach ML models to read charts made for humans, given the potential applications and contributions of such models to the HCI community. Among the main advantages of having ML interpreting charts are: the ability to quickly test our visualizations and interaction mechanisms with less reliance on human users, and the possibility to automatically generate visualizations based on the *auto-user*'s perceptual knowledge.

However, there are many challenges that need to be solved until such applications can become a reality. How do we evaluate the ML model, and assess how similar to the human user it is? One possible line of research is the use of model explanations, such as Anchors [15], or importance/activation maps [2], to investigate the reasoning behind the *auto-user*'s decisions. Another possibility is to add constraints to the *auto-user*, in order to to make it perform more similarly to a real human user. For example, since humans cannot discern more than 12 categories using a color hue encoding [11], we could replicate this behavior in the *auto-user* by reducing the number of colors in the chart, using a process known as color quantization.

In conclusion, adding ML to the HCI development loop might bring many benefits to the field. While there are challenges that we have to solve before this is possible, we feel it is important to start the discussion so that we can identify, as a community, what are the benefits and shortcomings of using such methodology, and what are the next steps to make it happen.

ACKNOWLEDGEMENTS

This project was partially supported by the NSF, the Moore-Sloan Data Science Environment, the DARPA D3M program and Labex DigiCosme. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of

DARPA. This work was performed while J. Freire and C. T. Silva were NYU Provost's Global Research Institute (GRI) fellows at NYU Paris.

REFERENCES

[1] S. Amershi, M. Cakmak, W. Knox, and T. Kulesza. 2014. Power to the People: The Role of Humans in Interactive Machine Learning. *AI Magazine* 35, 4 (Dec. 2014), 105.

[2] Z. Bylinskii, N. W. Kim, P O'Donovan, S. Alsheikh, S. Madan, H. Pfister, F. Durand, B. Russell, and A. Hertzmann. 2017. Learning Visual Importance for Graphic Designs and Data Visualizations. *ACM Symposium on User Interface Software and Technology - UIST '17* (2017), 57–69.

[3] W. S. Cleveland and R. McGill. 1985. Graphical perception and graphical methods for analyzing scientific data. *Science* 229, 4716 (1985), 828–833.

[4] Google AI. 2018. The What-If Tool: Code-Free Probing of Machine Learning Models. http://ai.googleblog.com/2018/09/the-what-if-tool-code-free-probing-of.html

[5] D. Haehn, J. Tompkin, and H. Pfister. 2018. Evaluating 'Graphical Perception' with CNNs. *IEEE Transactions on Visualization and Computer Graphics* (2018), 10.

[6] F. M Hohman, M. Kahng, R. Pienta, and D. H. Chau. 2018. Visual Analytics in Deep Learning: An Interrogative Survey for the Next Frontiers. *IEEE Transactions on Visualization and Computer Graphics* (2018).

[7] K. Kafle, B. Price, S. Cohen, and C. Kanan. 2018. DVQA: Understanding Data Visualizations via Question Answering. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, Salt Lake City, UT, 5648–5656.

[8] S. E. Kahou, V. Michalski, A. Atkinson, A. Kadar, A. Trischler, and Y. Bengio. 2017. FigureQA: An Annotated Figure Dataset for Visual Reasoning. *arXiv:1710.07300 [cs]* (Oct. 2017). arXiv: 1710.07300.

[9] Y. Lu, R. Garcia, B. Hansen, M. Gleicher, and R. Maciejewski. 2017. The State-of-the-Art in Predictive Visual Analytics. *Computer Graphics Forum* 36, 3 (June 2017), 539–562.

[10] D. Moritz, C. Wang, G. L. Nelson, H. Lin, A. M. Smith, B. Howe, and J. Heer. 2019. Formalizing visualization design knowledge as constraints: Actionable and extensible models in Draco. *IEEE transactions on visualization and computer graphics* 25, 1 (2019), 438–448.

[11] T. Munzner. 2014. *Visualization analysis and design.* CRC press.

[12] K. P. Murphy. 2012. *Machine learning: a probabilistic perspective.* MIT Press, Cambridge, MA.

[13] C. Plaisant. 2004. The challenge of information visualization evaluation. In *Proceedings of the working conference on Advanced visual interfaces*. ACM, 109–116.

[14] J. Poco and J. Heer. 2017. Reverse-Engineering Visualizations: Recovering Visual Encodings from Chart Images. In *Computer Graphics Forum*, Vol. 36. Wiley Online Library, 353–363.

[15] M. T. Ribeiro, S. Singh, and C. Guestrin. 2018. Anchors: High-Precision Model-Agnostic Explanations. In *AAAI*.

[16] M. Ricci, J. Kim, and T. Serre. 2018. Same-different problems strain convolutional neural networks. *arXiv:1802.03390 [cs, q-bio]* (Feb. 2018). arXiv: 1802.03390.

[17] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, and A. Bolton. 2017. Mastering the game of Go without human knowledge. *Nature* 550, 7676 (2017), 354.

[18] T. Weyand, I. Kostrikov, and J. Philbin. 2016. Planet-photo geolocation with convolutional neural networks. In *European Conference on Computer Vision*. Springer, 37–55.

[19] K. Wongsuphasawat, D. Smilkov, J. Wexler, J. Wilson, D. Mane, D. Fritz, D. Krishnan, F. B. Viegas, and M. Wattenberg. 2017. Visualizing Dataflow Graphs of Deep Learning Models in TensorFlow. *IEEE Transactions on Visualization and Computer Graphics* (2017).