The (un)predictability of Text-Based Processing in Machine Learning Art

Winnie Soon
Aarhus University
wsoon@cc.au.dk

Abstract

This article investigates the unpredictable vector of liveness within the context of machine learning art with a focus on text-based processing. [1] It is observed that there are similarities between generative art and machine learning art as both produce unpredictable results. According to Noah Wardrip-Fruin, the generative art form, such as *Loveletters* (1952), can be considered as a system that generates unpredictable outcomes. [2] *Loveletters*, allegedly the first digital literary work by computer scientist Christopher Strachey, is regarded as an 'unpredictable manifestation' of a system. [3] This system generates different variations of *love letters*, and such unpredictable manifestation is conditioned by two hidden elements: data and processes. The use of random algorithms plays an important role in generative art (Turing's random algorithm with its' random number generator was used in *Loveletters*) to produce autonomous and unpredictable outcomes. However, machine learning emphasizes 'predictive power,' in which prediction is produced through feeding in a large amount of training data. [4] Additionally, this kind of system employs predictive models and statistical algorithms to accomplish data processing and analysis. Machine Learning Art, such as text/novel generators, is claimed to be able to produce text with the similar writing style of the provided training corpus, but it also produces unpredictable text through setting different control parameters, such as number of epochs, amount of neural network layers and their hidden units, temperature and batch size.

This article is the result of the experiment of an open source machine learning library called ml5.js, which is built on top of TensorFlow.js, a Javascript framework, for training and deploying machine learning models. [5] ml5.js provides immediate access in the web browser to pretrained models for generating text. A Python training script employs the tensorflow library, which is used in the ml5.js environment to take in a large amount of text, and train a custom dataset as a pretrained model [6]. The study of the javascript libraries and the python script, with a specific focus on next character prediction and recurrent neural networks (RNN), unfolds the machine learning processes from data training to Long Short-Term Memory networks. [7][8] Building upon the notion of generativity, this article discusses the (un)predictable vector by examining the intertwining force between predictability and unpredictability that constitutes the liveness of text-based processing in machine learning art. [9][10][2] This paper argues that the (un)predictable vector of liveness helps to build an understanding of the relation between, but not in separation, training and execution processes, as well as the resultant actions that extend the aesthetic and live experience of machine learning art. The article contributes to the border understanding of generativity and liveness in machine learning art that employs generative models.

References

1. Winnie Soon, "Executing Liveness: An Examination of the Live Dimension of Code Inter-actions in Software (Art) Practice,"(Ph.D. diss., Aarhus University, 2016.)

2. Noah Wardrip-Fruin, "Digital Media Archaeology: Interpreting Computational Processes." In *Media Archaeology: Approaches, Applications, and Implications,* edited by Erkki Huhtamo & Jussi Parikka (Berkeley: University of California Press, 2011).
3. Noah Wardrip-Fruin, "Digital Media Archaeology: Interpreting Computational Processes" In *Media Archaeology: Approaches, Applications, and Implications,* edited by Erkki Huhtamo & Jussi Parikka (Berkeley: University of California Press, 2011), 306.
4. Adrian Mackenzie, "The Production of Prediction: What Does Machine Learning Want?" *European Journal of Cultural Studies* 18, nos. 4-5 (2013):429-445.
5. NYU. ITP "ml5js · Friendly Machine Learning for the Web." ml5js website. Accessed July 16, 2018. https://ml5js.org/.
6. NYU.ITP "Training a MSTM network and using the model in ml5js." ml5js github website. Accessed October 15, 2018. https://github.com/ml5js/training-lstm/.
7. Christopher Olah, "Understanding LSTM Networks (2015)." colah's blog. accessed July, 16, 2018. http://colah.github.io/posts/2015-08-Understanding-LSTMs/.
8. Andrej Karpathy, "The Unreasonable Effectiveness of Recurrent Neural Networks (2015)." Andrej Karpathy blog. accessed July, 16,2018.http://karpathy.github.io/2015/05/21/rnn-effectiveness/.
9. Philip Galenter, "What is Generative Art? Complexity Theory as a Context for Art Theory." (Paper based on a talk presented at GA2003–6th Generative Art Conference in Citeseer, 2003). https://www.philipgalanter.com/downloads/ga2003_paper.pdf
10. Philip Galenter, "Generative Art Theory." In *A Companion to Digital Art*, ed. Christiane Paul (Wiley-Blackwell, 2016).

Biography

Winnie Soon is an artist-researcher who is born in Hong Kong and is currently based in Denmark. Informed by the cultural, social and political context of technology, Winnie's work approach spans the fields of artistic practice, media art, software studies, cultural studies and computer science, examining the materiality of computational processes that underwrite our experiences and realities in digital culture via artistic and/or coding practice. Her works explore themes/concepts around digital culture, specifically concerning internet censorship, data circulation, real-time processing/liveness, and the culture of code practice, etc. Winnie's projects have been exhibited and presented internationally at museums, festivals, universities and conferences across Europe, Asia and America. Her current research focuses on exploratory and aesthetic programming, working on two books titled *Aesthetic Programming: A Handbook of Software Studies, or Software Studies for Dummies* (with Geoff Cox) and *Fix My Code* (with Cornelia Sollfrank). She is Assistant Professor at Aarhus University. More info: http://www.siusoon.net