Rapid prototyping: using Wizard of Oz to emulate machine learning features for interactive artistic applications

Alessandra Aleluia[1]*, Giordano Cabral[1]

[1]Mustic Research Group
Centro de Informática
Universidade Federal de Pernambuco (UFPE)

aaa6@cin.ufpe.br, grec@cin.ufpe.br

Abstract. *In this paper, we present using the Wizard of Oz method to rapid prototyping machine learning features in interactive artistic applications. Machine learning systems often require time and resources until they were available to be executed. But in the era of agile movement where we have to test as soon as possible and in artistic solutions whose scope initially perhaps be unclear, there is a need for a way to fast testing hypothesis. We also briefly described* Ama*, a system for adult ballet training at home, which served as a proof of concept of our strategy.*

1. Introduction

Machine learning is very often used in several applications, e.g. making predictions or to creating art. Although the design and implementation of systems that use it remain a contemporary challenge not only for developers but for designers too [1]. Machine learning classification techniques usually require a certain amount of data and time to train and tune the algorithms [2]. This scenario is considering they already had good datasets to use these techniques, which is not regularly the case.

Despite few datasets available for specif domains, artists so much as developers could not knowing in advance the final goal of the application and every step needed to get there. In recent years, with the ascending of agile movement and lean innovation, there's a need to rapidly test and evaluate solutions in the early stages of development to avoid wasting resources. It's expensive waiting for the full development of machine learning algorithms to get some feedback and we need to think in ways, with help of design techniques, to overcome these issues, for example, with participatory design [3]. Although there's no consensus about the prototyping method for machine learning experiences [4], in the case of interactive artistic applications we are investigating the use of the Wizard of Oz strategy to emulate machine learning features.

This paper is organized as follows: In Section 2 we discussed the ascending use of machine learning in artistic solutions, the issues around the use, the culture of lean innovation, and the Wizard of Oz method. Section 3 presents where we applied our approach, Section 4 was to describes it and in Section 5 we discuss what we learned through the process. We end this paper by outlining some limitations of the proposed approach and the hopes for the future.

*Supported by CNPq.

2. Background

Artists could benefit from the use of machine learning in their works using it in innovative ways. Now we have commercial software in the field of music AI to help in producing music without the need for music technologists to code. However a survey showed these people were a little skeptical about the current state of music AI but positive looking forward to the future of this field and many of them had built their own solutions using Python, SuperCollider, and Max/MSP [5]. Immersive experiences in mixed-media installations like *Where Is The Quiet?* [6] use machine learning to motivate individuals to relax, evoking meditative states of consciousness, and this kind of environment could support humans and provide new ways to appreciate art. Machine learning techniques are used in pose and gesture recognition, e.g., to support ballet training [7].

Machine learning algorithms improve through experience [8]. In general, they follow a well-established procedure: given a dataset, features were selected and extracted, these inputs were sent to the algorithm to train, some adjustments were made and the algorithm was ready to be used. It doesn't matter if it is a supervised, semi-supervised or non-supervised algorithm, these steps are almost the same as illustrated in Figure 1.

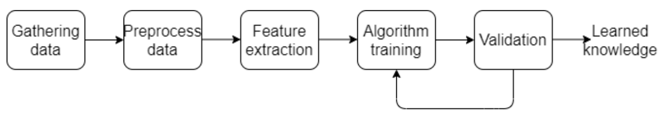

Figure 1: Machine learning process

In machine learning, it's fundamental to have a dataset to train your algorithms. Although we have some datasets available like ImageNet [9] for the computer vision field and efforts to develop platforms of openly available audio datasets to music computing like Freesound [10], they are not domain-specific and don't work for every application. Even though there are a few efforts to create APIs to support interactive machine learning, it doesn't solve the problem of not having time to create a very robust dataset to train machine learning algorithms. Research to explore the machine learning API design facilitating the usage for every kind of user didn't point out the problem of creating the dataset in a rapid prototyping approach [11].

Feature extraction is not a new problem and have been investigating for several years [12] and is complex and time-demanding [13], [14], [15]. We have different types of features, to know - low, mid, and high-level features. In Camurri et al.'s work, we can see the mapping of low and mid-level features into sonification and they propose other high-level features to enhance the expressiveness [16].

However not every artist knows how to code and even she/he does, machine learning algorithms are very often time-consuming in terms of training classifiers and creating datasets. A priori, it's common to the artist doesn't have a clear idea of what she is going to do, it is an exploratory approach. Buxton (1997) [17] states that the artistic spec on design level is the hardest and the most important one, this implies a need for rapid prototyping to not waste an amount of time on something already difficult. The needs of lean innovation approach are required in this era of Industry 4.0 with artificial intelligence technologies and connected devices and it will remain relevant for the next era, Industry 5.0, of close cooperation between humans and machines with cognitive computing [18]. There are some clues that the Wizard of Oz might be a good method for prototyping machine learning experiences, however, it was not tested with users yet [19].

Wizard of Oz consists of a low-fidelity prototyping method where the human, known by the wizard, simulates the software's response and it is used in prototyping AI systems too [20]. In this method, the users believe that they acting with a real working system, but it's only a human controlling it. Back in time, the use of Wizard of Oz simulations is well known in human-computer interaction in natural language. It was not easy to acquire high-quality empirical data in some cases, then the Wizard of Oz is a way of dealing with it. One of the uses of this strategy is to overcome the computer's limitations. Humans are not machines so they might have issues regarding timing and consistency. Therefore, to make this difference small as possible environments could be built for conducting experiments [21].

3. Ama: adult ballet training at home

We proposed *Ama: treinos feitos para você* to help ballet dancers to train at home without a teacher. The goal is not to replace the teacher, Ama's main objective is to assist adult students in practicing out of the dance studio. Taking into account their own needs, providing automatic feedback.

The process to create Ama was very interactive and we made use of participatory design[20], always talking with ballet dancers, designers, and programmers. More than 25 dancers participated in the process and it is still a work in progress. We needed to understand our target audience, integrated them into the design process, and figure out how to build a good solution together without wasting much time on technical issues. We did not know all the requisites that our system requires, we can only perceive it when we test it, show it to the users, and receive their feedback - emphasizing the need for a rapid prototyping approach.

Ama is a very complex system that requires many steps until it's fully developed and ready to be used. At first glance, we mapped seven main modules, to be said: tracking, recognition, recommendation of ballet steps, calibration, evaluation, feedback, and demonstration. Every module has its own challenges to be implemented and tested. Tracking and recognition, as a major important part, could be simulated by our proposed approach and we could be further in our research without spending time creating the dataset and training the machine learning to recognize the skeleton and the joints status.

The prototype of our system, using Wizard of Oz to emulate the machine learning features and an expert system to provide automatic real-time verbal feedback, was tested by 19 adult ballet dancers. They had two sessions one week apart and each person performed five ballet steps each week. Participants were asked to answer a questionnaire with five questions on the Likert scale to evaluate the received feedbacks and after that, they were interviewed. The interviews were transcribed and analyzed later. The results indicate the system's feedback was relevant (considering aspects like information, reinforcement, motivation, analysis, and critical thinking which are the functions of feedback [22]) and participants are excited to use it in real life.

4. Wizard of Oz applied in feature extraction

As a result of time restrictions and not knowing the full ultimate goal of the experiment in advance, being an exploratory work, the need for rapid testing our hypothesis, as well as a proof of concept, made us use the Wizard of Oz's approach to emulate the machine learning features which will be required to our system. Our case is the adult ballet scenario that hasn't, in our prior knowledge, no specific domain dataset.

As we stated earlier, it's time-consuming to create a specific domain dataset from scratch. Due to the coronavirus pandemic (COVID-19) in 2020, Brazilian studios of dance were closed and we couldn't have easy access to the ballet dancers - which made it even difficult to create a dataset. In gesture recognition the goal is, given a person often represented by a skeleton, to identify the action made. We were inspired by skeleton's representation made by Kinect [23] and Android ML Kit[1].

The Wizard of Oz strategy is commonly applied to simulate the whole system by a person. But, in our case, the use of this technique was only a part of the full prototype as we can see in Figure 2. Applying the Wizard of Oz technique is an often approach in design, especially in the conception of the prototype. However, here we are suggesting the use of a very technical part usually designated to programmers or people with consolidated coding skills.

In order to provide relevant real-time feedback, specific features needed to be extracted when the ballet

[1] https://developers.google.com/ml-kit/vision/pose-detection

dancer was performing. We developed an expert system based on rules to provide corrective and value verbal feedback. Features extracted using Wizard of Oz were the input of this system which gives the automatic feedback. Using this approach we can, with much less effort, rapidly prototype, test our hypothesis, and validate the system.

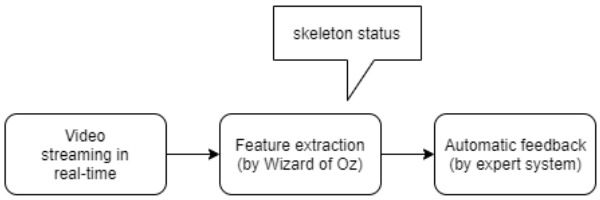

Figure 2: Our use of Wizard of Oz

Our wizard was responsible to enabled or disable the checkbox and click on the correct radio according to the execution of the step demonstrated. The checkboxes were: right step, right side, right arms, right head, stretched foot. The radios were: En dehors (en dedans, little open, open, very open), Posture (aligned, misaligned), Leg's extent (not applicable, low, mid, high), Elbows (not applicable, low, right, high) and Shoulders (tense, relax). They represented the features that we were willing machine learning extracts.

The application was in real-time, the wizard must send it many times to reflect the changes made by the dancer over the time of performance. Every feature was pre-defined and the evaluation was taught to the wizard so she could judge based on the same parameters all the executions. In Figure 3, there is the interface manipulated by the wizard while she is seeing the dancer's performance. Of course, because the feature extraction relies on the human operating the machine, feedback timing is influenced by the wizard's skills sending the status to the system.

Figure 3: Interface manipulated by human to emulate the feature extraction

5. Discussion

The use of machine learning is increasing in computing and in the intersection field of art and technology too. But we can put more effort to make a dialogue between the technical side and the design side, which is so important in experimental artistic solutions. Here we are proposing the use of a very useful design technique - the Wizard of Oz -, an old friend of us in the field of human-computer interaction to discover new horizons in the field of machine learning. However, we are not simulating the whole system as we have seen before, but the technique is playing an important role as a part of a major process.

Indeed in our case, the machine learning features that we are emulating were based on the current knowledge about human skeleton recognition and tracking. We did not mean to propose something far away from the reality or the current state of the art - which would not necessarily mean a problem if it were. So we spent a little bit of time thinking and organizing what informations we wanted to extract from the skeleton and formalizing them to be understandable for the wizard.

The system, an expert system based on rules, did not take too long to be implemented. We used Python language with the help of PySimpleGUI[2] to create the interface operated by the wizard. Most of the current machine learning applications are implemented in python, which is why we choose it - to make it easier in the future to integrate our solution with a real machine learning algorithm to extract the skeleton features. Our goal was to test our hypotheses - mapping some skeleton features to give useful real-time verbal feedback - and this approach was the way to rapidly prototype the system without the creation and training of a domain-specific dataset.

Time restriction and the difficulty to create a dataset from scratch were critical aspects that lead us to choose this approach, but they were not the only ones. We can go further. We need to think of new ways to approximate artists - and even programmers - who do not know how to build a very robust and reliable machine learning algorithm. Show them that we have ways to collaborate, and in the very beginning, there is no need to worry about such technical issues. We can go beyond it and test and retest our artistic solutions. Solutions with machine learning sometimes are seeing as a big technical challenge, and in fact, they are. But we can not let it stop us. We can overcome this challenge in the early steps to make us free to try

[2]https://pysimplegui.readthedocs.io/en/latest/

and test and make mistakes in the beginning.

6. Conclusion, limitations and future work

We presented a fast, easy and innovative way to prototype systems that made use of machine learning. This strategy could reduce the technology gap for people who do not want to waste much time implementing machine learning systems and focus more on the result instead of the technique. But, in terms of this work's limitations, maybe the person who emulates the machine learning features might be a little greedy and the feature extraction simulated were not in the granularity desired when it comes to real implementations - which is not necessarily a bad problem because it may push through the limits of the state of the art of machine learning techniques.

It's important to recall many questions are still open and many research opportunities are available. This paper only states the first step of a whole recognition system. We intend to implement machine learning algorithms that our Wizard of Oz strategy emulated. If we could track accurately human skeleton joints using an RGB camera, we will use it to train these algorithms and provide useful feedback. The correct tracking can be suitable for others problems that deal with gesture recognition too. We are looking forward to seeing other researchers in this diverse field making use of this method to overcome machine learning issues and making a lot of contributions in artistic areas.

References

[1] Qian Yang, John Zimmerman, Aaron Steinfeld, and Anthony Tomasic. Planning adaptive mobile experiences when wireframing. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*, pages 565–576, 2016.

[2] Aized Amin Soofi and Arshad Awan. Classification techniques in machine learning: applications and issues. *Journal of Basic and Applied Sciences*, 13:459–465, 2017.

[3] Tone Bratteteig and Guri Verne. Does ai make pd obsolete? exploring challenges from artificial intelligence to participatory design. In *Proceedings of the 15th Participatory Design Conference: Short Papers, Situated Actions, Workshops and Tutorial-Volume 2*, pages 1–5, 2018.

[4] Graham Dove, Kim Halskov, Jodi Forlizzi, and John Zimmerman. Ux design innovation: Challenges for working with machine learning as a design material. In *Proceedings of the 2017 chi conference on human factors in computing systems*, pages 278–288, 2017.

[5] Shelly Knotts and Nick Collins. A survey on the uptake of music ai software. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 499–504, 2020.

[6] McLean J Macionis and Ajay Kapur. Where is the quiet: Immersive experience design using the brain, mechatronics, and machine learning. In *NIME*, pages 335–338, 2019.

[7] Margaux Fourie and Dustin van der Haar. Ballet pose recognition: A bag-of-words support vector machine model for the dance training environment. In *International Conference on Information Science and Applications*, pages 317–325. Springer, 2018.

[8] Tom M Mitchell et al. Machine learning. 1997.

[9] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255. Ieee, 2009.

[10] Eduardo Fonseca, Jordi Pons Puig, Xavier Favory, Frederic Font Corbera, Dmitry Bogdanov, Andres Ferraro, Sergio Oramas, Alastair Porter, and Xavier Serra. Freesound datasets: a platform for the creation of open audio datasets. In *Hu X, Cunningham SJ, Turnbull D, Duan Z, editors. Proceedings of the 18th ISMIR Conference; 2017 oct 23-27; Suzhou, China.[Canada]: International Society for Music Information Retrieval; 2017. p. 486-93.* International Society for Music Information Retrieval (ISMIR), 2017.

[11] Francisco Bernardo, Michael Zbyszynski, Mick Grierson, and Rebecca Fiebrink. Designing and evaluating the usability of a machine learning api for rapid prototyping music technology. *Frontiers in Artificial Intelligence*, 3(13):1–18, 2020.

[12] D.M Gavrila. The visual analysis of human movement: A survey. *Computer Vision and Image Understanding*, 73(1):82–98, 1999.

[13] Basura Fernando, Elisa Fromont, and Tinne Tuytelaars. Mining mid-level features for image classification. *International Journal of Computer Vision*, 108(3):186–203, 2014.

[14] Vassileios Balntas, Andreas Doumanoglou, Caner Sahin, Juil Sock, Rigas Kouskouridas, and Tae-Kyun Kim. Pose guided rgbd feature learning for 3d object pose estimation. In *Proceedings of the IEEE international conference on computer vision*, pages 3856–3864, 2017.

[15] Haikuan Wang, Feixiang Zhou, Wenju Zhou, and Ling Chen. Human pose recognition based on depth image multifeature fusion. *Complexity*, 2018, 2018.

[16] Antonio Camurri, Gualtiero Volpe, Stefano Piana, Maurizio Mancini, Radoslaw Niewiadomski, Nicola Ferrari, and Corrado Canepa. The dancer in the eye: towards a multilayered computational framework of qualities in movement. In *Proceedings of the 3rd International Symposium on Movement and Computing*, pages 1–7, 2016.

[17] Bill Buxton. Artists and the art of the luthier. *ACM SIGGRAPH Computer Graphics*, 31(1):10–11, 1997.

[18] Banu Ozkeser. Lean innovation approach in industry 5.0. *The Eurasia Proceedings of Science Technology Engineering and Mathematics*, (2):422–428, 2018.

[19] Jacob T Browne. Wizard of oz prototyping for machine learning experiences. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–6, 2019.

[20] Yvonne Rogers, Helen Sharp, and Jenny Preece. *Interaction design: beyond human-computer interaction*. John Wiley & Sons, 2011.

[21] Nils Dahlbäck, Arne Jönsson, and Lars Ahrenberg. Wizard of oz studies—why and how. *Knowledge-based systems*, 6(4):258–266, 1993.

[22] Elizabeth Gibbons. *Teaching dance: The spectrum of styles*. AuthorHouse, 2007.

[23] Pichao Wang, Wanqing Li, Philip Ogunbona, Zhimin Gao, and Hanling Zhang. Mining mid-level features for action recognition based on effective skeleton representation. In *2014 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, pages 1–8. IEEE, 2014.