Newcastle University

Visualization to explain different aspects of Machine Learning

This proposal will explain the aims and objects of this project and the work process I will undertake in order to achieve them. It will briefly summaries my motivation behind choosing this as my dissertation project and what problems I aim to solve.

Motivation and rationale

Machine learning, being a core sub area of artificial intelligence allows computers to effectuate tasks which are practically impossible to code. The machines are trained to respond to tasks and instructions as a result of the previous input data that teaches the machines to eventually make independent decisions. In today's lifestyle of advanced technology, people are finding solutions to even the most complicated tasks with the help of AI and machines. "…our daily lives are becoming more influenced by "machine-generated" insights. The web pages you browse seem to know the next questions you would like to ask…." [1]
Thus, it becomes a necessity for people to understand the working of these complex mechanism of machine learning so that they can understand how they work to safeguard their own security and make their daily tasks easier.

The Problem

Besides the seeming necessity and functionality of machine learning, it becomes quite important that people are aware of the existing problem of bias in these algorithms. Bias is an evolving field of research where the machine learning algorithms are partial either towards it's developer or the input data.
The 'decision-making' by the algorithms works based on the previous training data that had been entered in it and uses this data to calculate the future outputs. This is rightly stated in Nicole Shawdowen's article, "It is important to note that machine learning algorithms *rely* on bias—statistical bias." [1]
Thus, as essential as the entire phenomenon of learning the process of machine learning might be, it is also crucial to notice that as the output is a result of bias, the algorithms cannot be trusted entirely.

Aims and Objectives

This project aims at educating people about machine learning and the risks thereof. This will be achieved by developing a visualization in the form of a website that allows the user to input a specific data (training data) and get the required output along with a step by step process of how the machine achieved the results. The steps would be an underlying mechanic of how a machine learning algorithm works. It would also allow people to understand the different aspects of machine learning. As the target audience of the project is the general public or non- programmers, the project aims at a simple web page that can be easily accessible by everyone without any requirement of login, download or any other data. Simplicity and accessibility will be a priority while designing this visualization so that people can learn about the procedure of decision making in a detailed but uncomplicated way.

Objectives

1) Research and identify three different aspects of machine learning that can be introduced via the visualization. These aspects would be the selected on a priority basis keeping in mind all the basic concepts every individual should know about machine learning.
2) Design a visualization that explains the key steps of machine learning and decision by making a website using HTML5, JavaScript, and relevant libraries (e.g. processing or d3.js) The model should focus on simplicity and should be approachable to everyone.
3) Authorize that the chosen algorithm doesn't support as much bias in decision making to avoid inaccuracy in the results.
4) Train the model rigorously to produce results on basis of training data. This would ensure that the users understand the underlying concept of machine learning.
5) Use pattern recognition to sort and distinguish data from each other. Strong patterns of previous data will likely generalize to make accurate predictions in future data. Decisions would also be made on the characteristics of pixels, color, and size of images.
6) The algorithm and the model should be robust enough to cope with imperfect data and extract repetitiveness to produce almost correct outputs. This objective raises some risk of inaccuracy along the way but should overcome by training and improving the algorithm.
7) Testing the mentioned visualization on some people and seeing if it works, if not modifying accordingly. This would be a final test for accuracy of the model.

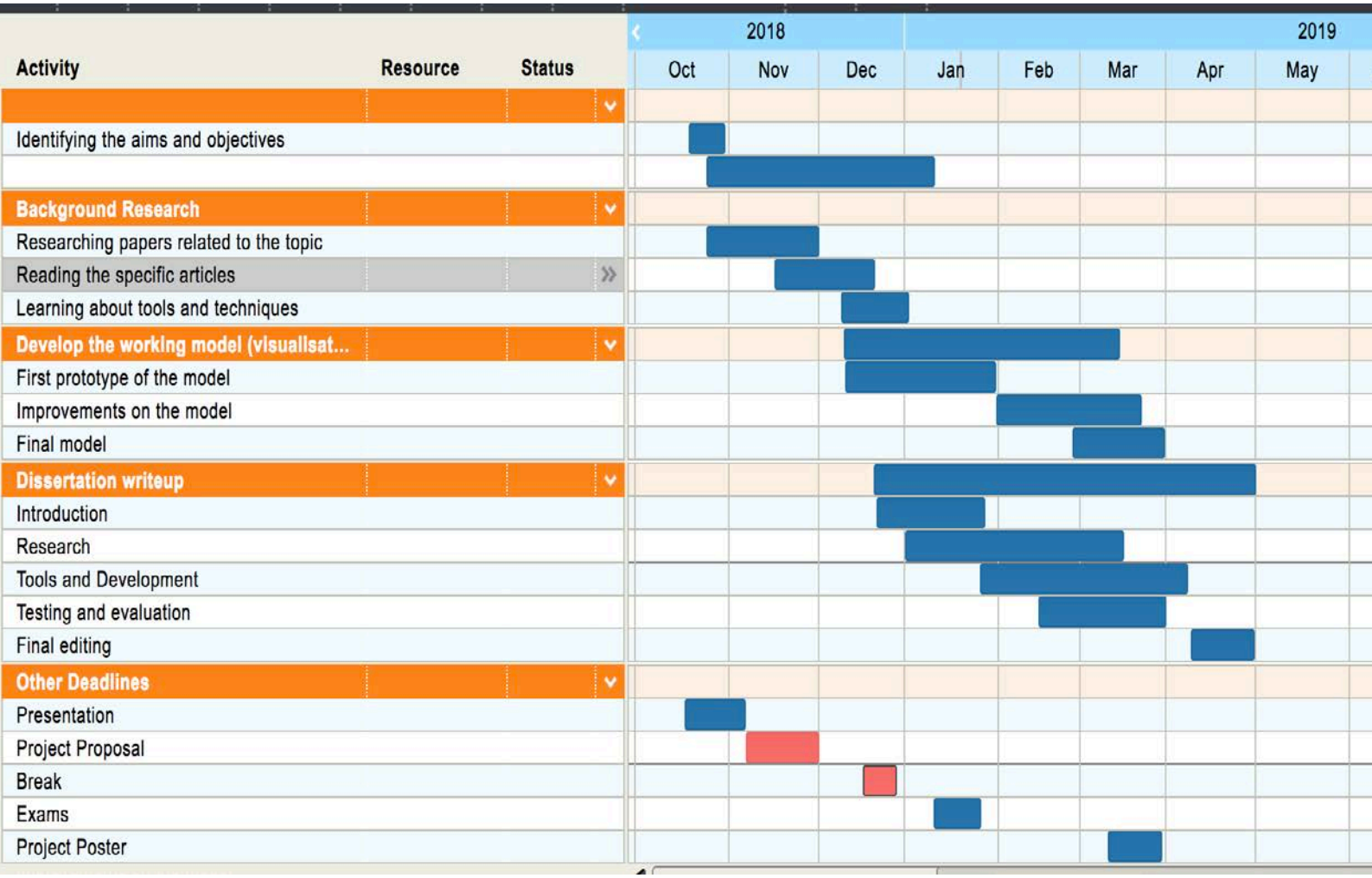
Background Research

Paper	**Summary and Relevance**
Hutton, T. (2007). *The Organic Builder: A Public Experiment in Artificial Chemistries and Self-Replication*.	This paper talks about a java applet where the rules of artificial chemistry can be experimented by non-programmers. This user-friendly interface is the closest example of my project. Where my project aims at creating a model that educates people about the different aspects of machine learning, Organic Builder is a platform which does the same in the field of chemistry. The author describes the design and the aims of the interface very systematically and lists how he has decided to achieve his goals by setting interactive challenges and experiments for users to play with and learn from the model.
Duygulu P., Barnard K., de Freitas J.F.G., Forsyth D.A. (2002) *Object*	This paper explains about a model of object recognition as machine translation which is one of the most similar objectives of my project. The author talks about how the model predicts the underlying concept of the input data and produces result by mapping images and extensive training. One of the few examples of

Recognition as Machine Translation: Learning a Lexicon for a Fixed Image Vocabulary.]	my project explains how my model will produce results on the basis of training data to get maximum accuracy eventually.
Shadowen, N. (2018). How to Prevent Bias in Machine Learning. [Blog]	This articles discusses the human perspective regarding machine learning along with three major topics- bias in machine learning, trusting the machines and the impact of bias in machine learning. This gives an insight of the power of machine learning in today's era along with an equal power of negative effect a bias algorithm can have on people's life. I learnt how to explain a particular topic from this articles and it also motivated me to consider things from a more social perspective.
Witten, I., Frank, E., Hall, M. and Pal, C. (2017). *Data Mining: Practical Machine Learning Tools and Techniques*. [	This book describes how tools and techniques of machine learning are used in practical data mining to find and describe structural patterns in data. As the author talks and explains about the different machine learning methods, it gave me an idea of the relevant tools and approach I should use for my project. The work introduces ways which are designed to explain the basic concepts. The book is represented by an illustration written in JAVA which explains all the contents of the book itself. Thus, as the aim of the book is similar to mine along with the visualization created to explain the book, it helped me to create a clear picture of what areas my project should include
Luo, G. Health Inf Sci Syst (2016) 4: 2.	This paper presents the first complete method for automatic explanation results for any machine learning model without degrading accuracy that make decisions on prediction. The key concept behind the model was to make two separate models, one based solely on predictions while the other would explain the results produced by the first model based on predefined rules.

Diagrammatic work plan

Fig 1.0



Explanation of Work Plan

I set an approximate working plan for the project during the first two weeks of the semester so that I could plan my work according to the upcoming assignments. I have always preferred working by setting soft as well as hard deadlines for my tasks so that I always have enough time for self-assessment and improvement in the end. I completed the important task of deciding on tools and technical areas for my project in the beginning so that even if a new software is required for the project, I have enough time to learn and get comfortable with using it.
The major problem with this project was that even though I was passionate about the topic, I didn't have any background knowledge of machine learning which would always arise the risk of either things not going according to the plan or me not being able to complete things on deadlines.

I have decided to do most of my technical work during the holidays, when I would have the most time so that even if my model doesn't work accurately, I could have enough time to make improvements. Until now I have researched articles and papers related to my project so that I can educate myself on the topics I want my website to include. As machine learning is a growing topic it is very important for me to thorough myself with all the researches. Thus, research is something that will be done over the entire period.
I have chosen the method of agile project management for this dissertation so that I can always go back to the previous tasks and make changes thereof. My project will be based on the critique made by my supervisor which will allow me to make changes along the completion of the project. I have been practicing HTML and JavaScript as these languages will be the platform of my technical work
The writing part of my dissertation will also start from December so that I have enough time to explain all topics extensively and also have enough time in the end to edit the final work.

References

[1] Shadowen, N. (2018). How to Prevent Bias in Machine Learning. [Blog] Available at: https://becominghuman.ai/how-to-prevent-bias-in-machine-learning-fbd9adf1198 [Accessed 29 Nov. 2018].

[2] Witten, I., Frank, E., Hall, M. and Pal, C. (2017). *Data Mining: Practical Machine Learning Tools and Techniques*. [online] Available at: http://thuvien.thanglong.edu.vn:8081/dspace/bitstream/DHTL_123456789/4050/1/Data%20mining-1.pdf [Accessed 29 Nov. 2018].

[3] Grgić-Hlača, N., Zafar, M., P. Gummadi, K. and Weller, A. (2017). *On Fairness, Diversity and Randomness in Algorithmic Decision Making*. [online] Available at: https://arxiv.org/pdf/1706.10208.pdf [Accessed 6 Nov. 2018].

[4] Amir Sani. Machine Learning for Decision Making. Machine Learning [stat.ML]. Université de Lille 1, 2015. English.

[5] Duygulu P., Barnard K., de Freitas J.F.G., Forsyth D.A. (2002) *Object Recognition as Machine Translation:*

[6] *Learning a Lexicon for a Fixed Image Vocabulary*.] Available at: https://link.springer.com/chapter/10.1007/3-540-47979-1_7

[7] Hutton, T. (2007). *The Organic Builder: A Public Experiment in Artificial Chemistries and Self-Replication*. [online] Available at: http://www.sq3.org.uk/papers/ob_report.pdf [Accessed 29 Nov. 2018].

[8] N.V., T. (2018). *Tom's Planner: Online Gantt Chart - Project Planning software*. [online] Tomsplanner.com. Available at: https://www.tomsplanner.com/ [Accessed 30 Nov. 2018].