Statement of Purpose

Sahil Pattni

(sp72@hw.ac.uk)

Few topics have fostered such a sense of curiosity, wonder, and interest in me as the broad field of computer science has. The process of seemingly arbitrary words on a screen translating to logic amazed me when I first discovered it, and it still does as I conclude my undergraduate degree. It is for this reason that I wish to pursue a Master of Computer Science at the University of Ottowa. My interests include machine learning, data structures, and natural language processing. After obtaining my Masters, I wish to further pursue a career in data science and/or software engineering.

Under the guidance of **Professor Neamat El Gayar**, I am working on my undergraduate dissertation, where I aim to derive multi-element association rules from a transactional database, as a complement to the established Apriori algorithm [1]. My goal is to do this by representing the data as an undirected graph where the products are nodes and the edges are the association scores between them. From this graph, I would extract a minimum spanning tree (MST) such that only the strongest associations remain, and identify distinct segments from the MST using Markov Clustering. From each cluster, I then plan to derive association rules from the connections within them. I am building on the work of **Marucio A. Valle et al** [2], whose paper proposed the initial framework for single element association rule derivation from MSTs. This dissertation has allowed me to explore my interest in data structures and graph theory, presenting its fair share of challenges along the way, all of which I enjoyed solving. For example, transforming the database of 24 million entries to my requirements had a complexity of $O(n^2)$, making it computationally infeasible if carried out in a linear fashion. I was able to use the knowledge I gained from my undergraduate courses such as Operating Systems and Concurrency to develop an efficient multi-core parallelized solution.

In the four years it took me to achieve a BSc. in Computer Science with Honors at Heriot-Watt University, I have acquired an understanding of this scientific discipline that has only fueled my curiosity and interest further. I have worked on several interesting projects during these four years, but one project worth mentioning is the coursework for my Biologically Inspired Computation course. For this project, I wrote a Particle-Swarm optimized neural network completely from scratch without the support of libraries such as Tensorflow and Scikit-Learn, resulting in me understanding how neural networks work at a fundamental level, and in doing so I realized the power to solve arbitrarily complex problems that machine learning offers. Another project worthy of mention is the coursework for my Applied Text Analytics course, where I was tasked with collecting Twitter data (tweets) and classifying their sentiment. I decided to use a Long Short Term Memory Recurrent Neural Network to process the text as a sequence and classify the tweets on a five-point sentiment scale. For our Software Engineering and Professional Development course, I led a team of five people to develop a smart home automation system for the course's year-long group project, which required us to develop a product in an emulated industrial environment, performing tasks such as identifying and determining the requirements; producing a comprehensive risk analysis; forecasting operational costs and budgets; conducting usability studies with prototypes, and finally delivering the finished product. This project showed me the value of leadership and that there was more to the software development cycle than just the technical phases. I believe a Masters program would provide the right environment to allow me to take advantage of more leadership opportunities.

Outside the academic environment, my passion for computer science has driven me to work on several projects on the side; most recently, I worked with two colleagues to develop a cross-platform mobile application for a hackathon that allows non-profits to crowd-source their funding. I am also currently developing an automated cryptocurrency trading system that uses machine learning[1] to determine optimal market conditions in which to execute trades. Laying the groundwork for the latter has also fostered an interest in stochastic systems in general, and how machine learning holds the potential to approximate them. I have also leveraged the knowledge I've acquired from my undergraduate degree in the family manufacturing business, using VBA and Python to automate and streamline most of the administrative work.

At the University of Ottawa, due to the plethora of available opportunities, I believe that I will be able to gain higher exposure to computer science and with it, a deeper understanding of the various problems in the field. Being part of such an elite and diverse group would foster intellectual growth and enable me to work to my potential. It would be a privilege for me to spend a fruitful and rewarding time studying at the Ottowa campus.

References

[1] R. Agrawal and R. Srikant, "Fast algorithms for mining association rules," *Proceedings of the 20th International Conference on Very Large Databases*, pp. 487–489, 1994.

[2] M. A. Valle, G. A. Ruz, and R. Morrśs, "Market basket analysis: Complementing association rules with minimum spanning trees," *Expert Systems with Applications*, vol. 97, pp. 146–162, May 2018. DOI: `10.1016/j.eswa.2017.12.028`.

---

[1]I am still comparing different classifiers and architectures to determine which one would best be suited for this task.