Statement of Purpose

Sahil Pattni

(sp72@hw.ac.uk)

Applying to Master of Computer Science at University of Ottawa

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 still does as I conclude my undergraduate degree. It is for this reason that I wish to pursue a Master of Computer Science at University of Ottawa. My interests include machine learning, data structures and natural language processing. After obtaining my Master's, I wish to further pursue a career in data science and/or software engineering.

I am working on a my undergraduate dissertation under the guidance of **Prof. Neamat El Gayar**, the purpose of which is to derive multi-element association rules by representing - as a graph - the association between product nodes as an edge that connects them, extracting the minimum spanning tree such that only the strongest associations remain, and determining clusters from this resulting minimum spanning tree using Markov Clustering. From each cluster, multi-element rules can be obtained from the edges within the cluster. This is an alternative to the established Apriori algorithm [1]. I am building on the work of **Marucio A. Valle et al** [2], who proposed the initial framework for single element association rule derivation from minimum spanning trees. This dissertation has allowed me to explore my interest in data structures and graph theory, but it has come with its share of technical challenges, all of which I enjoyed solving. Transforming the large dataset to my requirements computationally infeasible if performed linearly with a complexity of $O(n^2)$, due to the millions of entries in the dataset. I was able to use the knowledge I had gained in courses such as Operating Systems and Concurrency to develop an efficient multi-core, parallelized solution to meet these challenges.

In the four years it took me to achieve a Bachelor in 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 that stands out in my mind is the project 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, further fueling my desire to learn more about the field, and about computer science in general. Another project that I recall fondly is the coursework for my Applied Text Analytics course, where we were tasked with collecting Twitter data (tweets) and classifying their sentiment. I decided to use an Long Short Term Memory Recurrent Neural Network to classify the tweets on a five point sentiment scale, and this project my interest in natural language processing further [talk more about NLP]. For our Software Engineering and Professional Development course, I led a team of five to develop a smart home automation system for the course project, which required us to develop a product in an emulated industrial environment, performing tasks ranging from identifying and determining the requirements, performing a comprehensive risk analysis, forecasting operational costs and budgets, conducting usability studies with prototypes, and finally delivering the final product.

Towards the end of the second semester of my third academic year, the world was hit with the Covid pandemic, and as universities, schools and businesses were closed under a government-mandated lockdown, my university did not have enough time to transition to an online examination system. Therefore, for courses whose grades were primarily determined through exams, a passing grade (P) was awarded since the exams could not be administered. For courses that were marked through courseworks, the regular grading system was used as all courseworks had been submitted by the time the lockdown was put in effect.

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 when to execute trades. Laying the groundwork for this project has also fostered an interest in stochastic systems in general, and how machine learning holds the potential to approximate them. I have also applied the knowledge I've gained in the family manufacturing business, using VBA and Python to automate and streamline most of the administrative work.

At University of Ottawa, due to the plethora of available opportunities, I believe that I will be able to gain higher exposure 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 will be a privilege for me to spend a fruitful and rewarding time at the Ottawa 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