

Habeebullah Ebrahemi

May 25, 2021

Letter of Intent

I intend to explore Network Science with the main focus on Social Network, as a part of my Ph.D. program. I believe that Network Science is an interdisciplinary field that encompasses the structure of complex networks such as telecommunication networks, computer networks, biological networks, cognitive and semantic networks, and social networks.

Previously, I was under the supervision of Dr. Muhammad Abulaish (South Asian University, India) whose guiding me in my initial research, A Graph-Based Approach for Sentiment Analysis over Microblogging Data which is submitted successfully. My focus was on methods that seek to apply graph based theories and approaches to enhance the current work efficiency in terms of time and space and effectiveness regarding accuracy metric and address the new challenges raised by sentiment-aware applications, as compared to those that are already present in more classical manner.

I was involved with a machine learning team after my graduation in 2019, in NeoEpo company, where I joined working as a Data Scientist Software Developer hands on various machine learning projects that can address real world business problems.

Currently I am invited to extend the paper that we submitted to WI (web intelligence) conference.(Muhammad Abulaish, Mohammad Rahimi, Habeebullah Ebrahemi, and Amit Kumar Sah, 2019)

What is **Sentiment analysis**? It is the process of using tools and tricks such as natural language processing and text analysis to find out what other people think in order to determine whether their attitude towards a particular topic, product, public figure, political event, etc. is positive, negative or neutral (polarity). It is essential because, consequently, we extract valuable knowledge for different purposes. (El-Din & Husein, 2016)

There are other methods that were initially introduced such as the n-gram model and term vector language model which are cost-effective regarding efficiency and effectiveness and also they are highly language dependent. (Aisopos, Papadakis & Varvarigou, 2011).

Meanwhile, there are many recent works like analyzing and detection of bipolar phrases and conflict bipolar that can support only German, French and English languages, in contrast, the character n-gram graph that I work on is language independent. (Klenner, Amsler & Hollenstein, 2017).

I am inspired by Moreno who claimed in the social network analysis: methods and applications "before the advent of sociometry no one knew what the interpersonal structure of a group 'precisely' looked like". (Moreno, 1953).

In the light of available information, it is observed that the first academic studies measuring public opinion were carried out during and after world war II and their objective was mostly political in nature. As against this, sentiment analysis in the ten years between 2005 and 2016 has witnessed a 50-fold growth. (Mäntylä, Graziotin, Kuuttila, 2018). However, there is still a lot of scope for further research.

The model that I have worked on, presents a language independent of microblog's model and mainly targets Twitter data. Although it is interesting and works fine with each

tweet that is in corpus but my aim is to go further deep in complex structure of social network where data is diverse and also to handle unstructured data in the form of text, image, videos or lots of new behavioral indicators that are being used on Facebook and many other social networks.

The motivation to pursue research in this area is a content restriction, other social media services offer communication rich in multimedia data, i.e., audio, video, images, etc. Whereas Twitter gives a limit, for instance, each user can tweet only 140 characters which may not be sufficient. (Giannakopoulos & Palpanas 2010).

My interest in Network Science developed while I was pursuing my undergraduate and graduate's degree. During my bachelor course, Khatera Barekzai 's class" Sociology1 and Sociology2' ' was the first to introduce me to the humans' behavior in society from pre-birth time until he/she joins the everlasting world. Kazim Hooshmad's "Networking" where we followed Cisco company devices and configurations, advanced my knowledge of computer networks from small to enterprise level.

'Ali Sajjad Haris 's class "Java1 and Java2" helped develop my programming skills because of which I was able to work as the leader of a team consisting of five members. Our team was engaged in a very interesting task as we developed our first Android-based Application and presented it at the end of the sixth semester. This encouraged me to do my BSc final project, web-based system, 'LMS'(Library Management system) for our own university and I also wrote all the accompanying documentation successfully.

At M.Sc level, Dr. Muhammad Abulaish's class on "Data Mining" introduced me to literature and theories of knowledge discovery through data and trained me about different algorithms in clustering and classification. Rashma Rastogi's (nee Khemchandani) "Machine

learning” introduced me to the state of art algorithms in machine learning SVM and many more classification algorithms. These assets led me to choose ‘A Graph-Based Approach for Sentiment Analysis over Microblogging Data’ as my subject of research. This is one of the areas that help in a lot of decision making cases from traditional approaches, product reviews to modern approaches, i.e., social media text analysis.

During my summer vacation, I went through a purely practical python programming course” Python programming: a concise introduction “from Coursera The approach was to present an example followed by a small exercise where the learner tries something similar to solidify a concept and I think that is so helpful during my project implementation.

Recently my semester-wise project was about Comparative Study of HITS and PageRank Link based Ranking Algorithms that I delivered last semester. (Kumar & Munish, 2010) In addition to my studies, I got hands-on practice using powerful programming language python and most of our implementations were in Python and Matlab.

During my winter break, I was enrolled in a course in an emerging field, i.e. Big Data. The course, “Introduction to Big Data” enabled me to gather the fundamental terminologies in the subject and grasp the vital role of Big Data in business and industry. The three main sources of Big Data, i.e., People, Organization, and Sensors; and seven V's, i.e., Volume, Velocity, Variety, Variability, Veracity, Visualization, and Value that will help me to get a lot of ideas about my field.

Considering my research interest and background, I believe I am one of the suitable candidates for pursuing a Ph.D. course.

Since last year, I have been one of the interested people of the ICCR programs. This has enabled me to go through some impressive work done by different scholars and it also showed me clearly the beauty of diverse nations coming together and collaborating for the same objective, which is to make the world a better place. People from different parts of the world with colorful backgrounds exchange ideas and consequently help in the emergence and rapid development of new ideas.

After obtaining my Ph.D. degree, I see my career as a combination of academic (scholarly) work as a professor and an entrepreneur for the new startups. I believe that in these fields I can make the best contribution, utilizing the advantages of my diverse theoretical and practical background. And I hope you will give me the privilege to move towards these goals at your truly unique and esteemed institution.

Works Cited:

"Muhammad Abulaish, Mohammad Rahimi, Habeebullah Ebrahemi, and Amit Kumar Sah, SentiLangN: A Language-Neutral Graph-Based Approach for Sentiment Analysis in Microblogging Data, In Proceedings of the 18th IEEE/ WIC/ACM International Conference on Web Intelligence (WI), Thessaloniki, Greece, October 14-17, 2019, pp. 461-465."

Aisopos, F. & Papadakis G. & Varvarigou T. (2011), "Sentiment analysis of social media content using n-gram graphs." Proceedings of the 3rd ACM SIGMM International Workshop on Social Media - WSM '11, doi:10.1145/2072609.2072614.

Giannakopoulos, G. & Palpanas. T. (2010), Content and type as orthogonal modeling feature: a study on user interest awareness in entity subscription services, International Journal of Advances on Networks and Services, 1-3

Kumar, Munish, Comparing the Effectiveness of PAGERANK, HITS, and SNorm (p) Web Page Ranking Algorithms. (2010), International Journal of Computer Applications.11,1-4., doi:10.5120/1590-2130.

Klenner, M., Tron, S., Amsler, M., & Hollenstein, N. (2017). The detection and analysis of bi-polar phrases and polarity conflicts.

www.semanticscholar.org/paper/The-Detection-and-Analysis-of-Bi-polar-Phrases-and-Klenn-erTron/7a7fd48731ed64b97a527d82271192c20de0490a

Mäntylä.V., Graziotin D., & Kuutila, M. (2018), The evolution of sentiment analysis - a review of research topics, venues, and top cited papers.Computer Science Review,27,16-32. doi.org/10.1016/j.cosrev.2017.10.002.

Last, F. M. (Year Published) **Book**. City, State: Publisher

Wassernab, S. Faust, K.(1999) Social network analysis: methods and applications. Retrieved From

https://books.google.co.in/books?id=CAm2DplqRUIC&pg=PA12&lpg=PA12&dq=before+the+advent+of+sociometry+no+one+knew+what+the+interpersonal+structure+of+a+group+%27precisely%27+looked+like&source=bl&ots=HwJtzh_CRC&sig=ACfU3U0PYB9509KUeN6vASaSExmt0vMlVg&hl=en&sa=X&ved=2ahUKEwjfvob3u5bgAhVLPI8KHZxrA-AQ6AEwAXoECAUQAQ#v=onepage&q=moreno%20claimed&f=false