Prospectus

Developing Spam Detection and Prevention Schemes using Natural Language Processing

(Technical Report)

An Analysis into the Efficiency of Makerspaces

(STS Research Paper)

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Parv Ahuja

Fall, 2019

Department of Computer Science

Signed: ___Parv Ahuja_____

Approved: _____ Date _____

Professor Tsai-Hsuan Ku, Department of Engineering and Society

Approved: _____Yuan Tian_____ Date 12/5/2019

Assistant Professor Yuan Tian, Department of Computer Science

INTRODUCTION

In recent years, the emphasis on science, technology, engineering, and mathematics (STEM) within the education system has dramatically reshaped the landscape of teaching and changed the way in which students engage in learning. Within STEM alone, education has progressed to involve more independent learning and erudition through exploration. One recent example of learning through exploration comes from the Makerspace movement happening in the United States, where makerspaces are being created in the name of STEM in remarkable numbers (Halverson, 2014). But what even is a makerspace? From a technical perspective, a makerspace is nothing more than a collaborative environment equipped with fancy tools, like 3D-printers and CNC machines. The goal of these makerspaces is to harbor creativity and encourage STEM values through hands-on exploration (Avneet, Hynes 2018). The growing makerspace movements within the United States and globally have put makerspaces on a pedestal - but are they truly effective within a community? Are the educational and community benefits reaped from Makerspaces worth the resources? Through this thesis, I intend to form an argument regarding the utility of Makerspaces as a means of community engagement and as a learning tool within the educational system.

In the technical portion of the prospectus, I discuss my preliminary research into the field of Natural Language Processing (NLP) and explore the possible applications of NLP for detecting and preventing spam. After validating the problem statement, the report introduces an artificially intelligent chatbox capable of detecting spam behavior and collecting threat intelligence. Following this, the importance and criticality of NLP within the context of the chatbox are discussed and possible future applications are explored. Finally, an in depth review of the various models and techniques used to process sentences is conducted.

STS THESIS

As the Makerspace movement grows within the United States, and more cities begin to adopt Maker culture, the need to evaluate the efficacy of Makerspaces becomes pertinent. Moreover, it is paramount that there exists a way to evaluate Makerspaces that considers both the associated resources and benefits. For instance, when constructing a Makerspace for a smart city, it would be important to consider the utilisation of community resources and the ability to quantifiably evaluate the benefit (or lack thereof). For the most part, the impact these spaces currently have on the surrounding community and the educational system are assumed. In other words, makerspaces are being measured in terms of their potential for impact, not their actual impact (Cun, 2019). Furthermore, the resources and cost associated with building and maintaining a makerspace are often overlooked. As the team proposes a blueprint to implement mobile makerspaces in and around Charlottesville public schools, I plan to research and investigate the overall efficiency of makerspaces, both from a community and educational lens. This project views makerspace as a complex socio-technical system constituted by multiple social groups and artifacts. I will use SCOT to investigate the efficiency of makerspaces, through examining the interaction of the community members and students. Through this investigation and further research, I hope to

answer the question: Are Makerspaces an efficient tool both within a community and throughout the educational system?

In answering this question, it is important to define the criteria to which we will measure the impact of Makerspaces on the community and education system. To help better understand the impact of makerspaces on a community, I investigated literature pertaining to frameworks to help analyse the socio-technical system. Hira & Hynes (2018) discuss various frameworks that analyse the effect of makerspaces from a nontechnical perspective. In one framework, the authors discuss the relationship between people, means, and activities, and then generate the representation of those in the context of makerspaces. The framework creates the groundings to analyse makerspaces as a whole, and further deduce the specific impacts on various actors within the framework. Although no specific result was achieved from the framework, it sets up a good basis to conduct more research into community impacts of makerspaces. Another article, called *An assessment matrix for library makerspaces*, is extremely pertinent to both my individual research into educational impact and the team's blueprint to produce makerspaces within Charlottesville libraries. Aijuan Cun, Samuel Abramovich, and Jordan Smith, in response to the emergence of maker culture, develop an assessment matrix to help assess learning benefit from Makerspaces. After analysing different theoretical perspectives on assessment, like summative and formative assessment, the authors implement a research methodology to iteratively design the assessment process. Through observation of key values within library makerspaces, data collection, and feedback from librarians, the authors were able to produce a concrete procedure by which to evaluate library Makerspaces (Cun, 2019). The assessment matrix, although specific to library environments, defines useful criteria that can be incorporated into my research. More specifically, I plan to utilize the author's defined assessment procedures for common maker activities, like 3-D printing and Virtual Reality. Furthermore, I hope to follow a similar design methodology that the author's used to define the roles of patrons (librarians) to investigate the role of community and educational actors in a more general Makerspace.

STS FRAMEWORK AND METHOD

In order to conduct my own research, I plan on utilising the social construction of technology framework (SCOT). SCOT is a framework that recognises the social groups, conflicts, and existing artifacts and relates them together. Another core concept of SCOT is the idea of interpretive flexibility, that states that different social groups interpret the same technology in different ways, and each of these different understandings can introduce a new problem (Bijker, 2015). In the context of makerspaces, I have preliminarily defined the relevant social groups to be the outlying community members and the students/faculty within the education system. Both of these social groups interpret the costs and benefits of makerspaces differently, and in addressing those issues, I hope to gain insight into the relationship between makerspaces and these two groups. Overall, I intend to understand this difference in interpretation of Makerspaces and utilise it to analyse the efficacy of Makerspaces in regards to these groups. After mapping these relationships, I will use a holistic approach to impute a value on creating and maintaining a makerspace, and then compare this to my SCOT analysis to draw final conclusions.

In addition to literature review and preliminary SCOT analysis, I have also completed some interviews to gain a well-rounded global insight into the maker movement. The interviews were with a student from Tsinghua University in China, and the purpose was to provide me with basic insight into the societal impacts of the maker movement within China. These discussions were informative, and allowed me to gain a different perspective on the maker movement. Furthermore, through discussion of Makerspaces in a different country, I was able to incorporate new ideas and previously unthought of impacts into my research. The interviews proved useful in grounding myself into a different cultural norm, allowing for an unbiased and overall more extensive assessment of Makerspaces, both within the United States and globally.

In the future, I would like to focus on two main things in order to fully connect my research thus far to my research question. First, I would like to expand on the assessment matrix discussed before to be applicable to both the education system and community members. In other words, I would like to pick and choose the relevant artifacts of the assessment matrix and use it to evaluate the engagement within a community. With this, I would be able to properly quantify the effect of makerspaces on both the community and the educational system. Next, I would like to further investigate the cost of Makerspaces in the short-term and long-term. As I have not researched this yet, I plan on creating a holistic approach to imputing the value of makerspaces that could be expanded to analyse the cost, monetary or other, of any makerspace.

TECHNICAL REPORT

Every day, hundreds of thousands of individuals sit in front of their computers and write fake reviews for products based upon the tasks they receive from the underground market for the purposes of manipulating product ranking. Spam callers, review writers, and other miscreants communicate using underground crowdsourcing, where they utilize instant messaging (IM) to coordinate their attacks (Undisclosed, 2019). The threat intelligence provided by this IM communication is invaluable to understanding and mitigating fraud, but hard to systematically gather and analyse. In comes Aubrey, the first autonomous chatbox that actively collects intelligence through communication with real-world miscreants. The chatbox, Aubrey, poses as a miscreant looking for tasks, and utilizes the underground communication pipelines to engage in conversation and extract relevant intelligence. To achieve autonomous conversation, Aubrey is modeled as a finite state machine where states represent different possible stages in the conversation, and state transitions are performed based on the responses of the miscreant. In order to facilitate this automata, Aubrey must be able to deduce meaning and intent from miscreant messages and then utilize that information to generate a valid response (Undisclosed, 2019).

Natural Language Processing is the field of study regarding natural human language and computer science. More specifically, NLP involves how to program computers and software to analyse and process linguistic data. In the context of Aubrey, NLP is used to ensure smooth sentence transitions and a seamless conversation. First, NLP is used to translate a miscreant message into some recognisable state for Aubrey. To do this, the message is broken up into different words, each

of which becomes mapped to a higher-dimensional vector that represents the word's relationship to the entire sentence. Next, the vectors are processed and translated to specific intents based on a previously defined database of relationships. Finally, a classifier uses the derived intent of the message and produces the state probabilities, essentially mapping the message to a state within Aubrey's finite state machine (Undisclosed, 2019). Recall, that states within the FSM were designed to represent different possible stages within a conversation with a miscreant, and this process essentially maps a miscreant message to a particular stage in the conversation that Aubrey can recognise. After this, Aubrey can follow the state transitions and produce a response that reflects understanding of the miscreant message.

From this prior research, it becomes evident that Natural Language Processing has a significant role in autonomous conversation . Even outside of the scope of Aubrey and e-commerce fraud, NLP is a key factor in the future of spam detection and prevention. Autonomous conversation with spam callers, chatting with e-commerce miscreants, and detecting spam emails are all potential schemes that would require NLP. The remainder of the research assumes the use of NLP in spam detection, and refocuses on the specific NLP techniques that could enable these detection and prevention schemes.

In recent years, there has been tremendous progress made in the field of natural language processing, including more powerful models and increased relational data. One of the most basic frameworks within NLP is a transformer, which is a mechanism to gather relevant context of a given word, and then encode the context into a multi-dimensional rich vector (Rizvi, 2019). Before BERT, a majority of transformers in application were long-short-term-memory based models, like the one applied for Aubrey. LSTM transformers operate much in the way that humans do when processing sentences, in that encoding and decoding for context happens left-to-right in a linear fashion (Rizvi, 2019). Although this model was powerful in its own manner, it ran into issues when context was hidden within a sentence. For example, in the two sentences below, a LSTM transformer would have trouble deducing the different meanings of the word "bank" looking only left to right (bolded words provide context).

1. I need to go to the bank **to deposit some money**
2. **We went to the river** bank

BERT, an acronym for Bidirectional Encoder Representations from Transformers, tackles this problem by jointly conditioning on both the left and right context. The tremendous framework was developed by Google AI just this year, in 2019, and has already made groundbreaking advancements in NLP research. BERT was able to use a single model to achieve state of the art results on 11 individual NLP tasks, using only an unlabelled dataset. Through personal investigation, in a simple classification problem, I was able to utilise a pre-trained BERT model to achieve higher accuracy than my previous fully trained LSTM model. In other words, without any finetuning or training, BERT was able to achieve better results than the LSTM model that had been fully trained on the dataset.

For the remainder of my research, I plan to delve into BERT and develop a model that could surpass results achieved within the spam detection field. In Aubrey, for example, a BERT model

could replace the current LSTM model to reduce error in deducing meaning from miscreant sentences, overall enabling Aubrey to collect more threat intelligence. Specifically, I propose implementing BERT within Aubrey and examining the results and more research into spam-specific NLP tasks that could be better solved using BERT.

BIBLIOGRAPHY

Note: As part of the technical report, I reference an article that is still being refined and under review. This article is key to my research, but due to its nature as a work in progress and per the request of my technical advisor, I have not included the author or title of the article.

UVA Graduate Student (2019) Aubrey: The autonomous e-commerce chatbox. UVA handbook: 2019

(Undisclosed, 2019)

Hira, Avneet, and Morgan M Hynes. "People, Means, and Activities: A Conceptual Framework for Realizing the Educational Potential of Makerspaces." *Education Research International*, Hindawi, 3 June 2018, www.hindawi.com/journals/edri/2018/6923617/.

(Avneet, Hynes 2018)

Bijker, Wiebe E. "Technology, Social Construction Of." *International Encyclopedia of the Social & Behavioral Sciences*, 2015, pp. 135–140., doi:10.1016/b978-0-08-097086-8.85038-2.

(Bijker, 2015)

Cun, Aijuan, et al. "An Assessment Matrix for Library Makerspaces." *Library & Information Science Research*, vol. 41, no. 1, 2019, pp. 39–47., doi:10.1016/j.lisr.2019.02.008.

(Cun, 2019)

Erica Rosenfeld Halverson and Kimberly Sheridan (*2014*) The Maker Movement in Education. Harvard Educational Review: December 2014, Vol. 84, No. 4, pp. 495-504.

(Halverson, 2014)

Rizvi, Mohd Sanad Zaki. "Demystifying BERT: The Groundbreaking NLP Framework." *Analytics Vidhya*, 25 Sept. 2019, www.analyticsvidhya.com/blog/2019/09/demystifying-bert-groundbreaking-nlp-framework/.

(Rizvi, 2019)

Sheridan, Kimberly, et al. "Learning in the Making: A Comparative Case Study of Three Makerspaces." *Harvard Educational Review*, vol. 84, no. 4, 2014, pp. 505–531., doi:10.17763/haer.84.4.brr34733723j648u.

(Sheridan, 2014)