Prospectus

NLP Comparison between Movie Reviews from Critic' and Consumers

(Technical Report)

The Role of Technology in Facilitating Contemplative Collaboration

(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

Yuelan Sheng

Fall, 2019

Department of Computer Science

Signed: Yuelan Sheng

Approved: Tsai-hsuan Ku Date 12092019

Sharon Tsai-hsuan Ku, Department of Engineering and Society

Approved: Date 11/25/2019

Yangfeng Ji, Department of Computer Science

Introduction

Problem solving is an important mission for all the engineers, and it is even more desired for solving problems that have been ignored. There are all sorts of topics from every area that engineers have not been able to touch on and as technology advances and people's understandings of technoscience broadens, there are better approaches in solving these mysteries.

In my technical capstone project, I work on exploring the differences in professional critics' reviews and consumer reviews using natural language processing to conclude better marketing strategies in using the reviews and build a reliable model based on all gathered information. The previous research did not compare the two groups of reviews from text levels and focused more on other characteristics. I am tackling a problem from a new perspective and I am aiming to transform data into insights that can be applicable in marketing of movie industry and maybe even extend to film production.

The STS thesis focuses on the topic contemplative collaboration, which is fairly new and innovative. The introduction of the idea is based on the process of designing the UVA contemplative collaboration space. Contemplation, which defines to be an individual activity, is promoted to improve the collaboration space dynamics. In the team project, we will try to see what types of adjustment can be made to the space and particularly in my own thesis, I will focus on the technological adjustment and how the technology works with both contemplation and collaboration.

Technical Topic - NLP Comparison between Movie Reviews from Critic' and Consumers

What insights and marketing strategies can be found by comparing the actual texts between professional critics and normal consumers using natural language processing?

Upon entering a new era in which big data and algorithms will revolutionize people's lives, there are more and more studies applying machine learning techniques in developing better marketing strategies for different industries. In the movie industry, in order to improve the box office sales, researchers and companies have been studying the potential factors that affect the selling of the movies. The past studies have been mainly focused on the characteristics of the movies, such as genres and leading actors/actresses. Some of the studies also apply natural language processing (NLP) techniques in figuring out how the written reviews of the movies can affect the box office performance. In reality, there are two major groups of movie reviewers, the professional critics and regular consumers. These two groups are highly possible to have different perspective and emphasis on evaluating the movies, which ultimately create reviews of different effects. There are huge marketing potential in applying the differences of reviews since the reviews can be promoted in a certain way that influence people's decisions of going to a movie or not, and the information of the reviews can be used as a model for predictions of the box office of a certain movie.

In order to better understand the differences in movie reviews from the two distinct groups and relate the differences to effects on box office sales, sentiment analysis, topic modeling, and other statistical analysis need to be conducted on the dataset.

To collect the consumer review data, we scanned through more than 10,000 movies on IMDB website and scrapped the consumers reviews from IMDB. The reviews from professional critics are scrapped using top-10 film review websites: the New York Times, San Francisco Chronicle, Roger Ebert, The Onion AV Club, Slant Magazine, eFilmCritic, Blu-Ray, Austin Chronicle, Pop Matters, and Common Sense Media. In order to eliminate potential biases, I cleaned the dataset by eliminating movies that do not have more than three reviews from both the critics and the users.

Before any NLP techniques are done, it is important to perform some evidence-free analysis using statistical mechanics. I trained different regression models and correlation matrices on the statistics of the reviews with the box office sales to see the correlations among them. Then I want to learn about the sentiment of reviews. After performing sentiment analysis on the dataset, a general pattern of whether a certain group of reviewers, professional critics or normal consumers, usually review the movie more positively or negatively. Within each movie, it is also possible to generate the idea of what contributes to the variation of sentiments of the reviews, the group of reviewers, or the genre of movies. In order to figure out what specifically the movie reviews talk about, I am performing topic modeling on all the effective reviews and see what are the topics the people generally include when evaluating a certain movie. Furthermore, the movie reviews are divided into different genres and analyzed by topic modeling again. After I get the result of the topics, I will label them with Professor Natasha Foutz, a marketing professor from the McIntire School of Commerce. Using the topic lists of the reviews and relating them back to the original reviews, Professor Foutz and I will try to understand what are the reasonings behind the different topics the reviews talk about and develop some hypotheses for future marketing strategies in the movie industry.

One of the challenges I faced was the size of the data, which exceeds more than 1.4 million entries. The huge number makes using normal NLP packages hard and I overcome the problem by using OnlineLDA model. The most challenging part of the project is that human languages are ambiguous and humans make errors when typing words, which result in the incomprehensibility of results and potential errors in concluding the general patterns. To overcome this difficulty, I talked with Professor Yangfeng Ji and decided to test out different models and allow more outlying groups to address the possibility of ambiguousness. Moreover, the professional knowledge from Professor Foutz about marketing strategies will also help to make a more thorough conclusions.

The final result should conclude the differences of reviews from two groups generally speaking and genre specifically speaking, and be able to explain the possible reasons for these differences and apply the findings to future marketing strategies and regressional predictions of box office.

STS Thesis - The Role of Technology in Facilitating Contemplative Collaboration

Introduction

During a normal meditation session, people are required to put away their electronic devices in order to let their minds concentrate on the contemplative activities, and this has proven to be effective. However, for most people, after they are done with the contemplation practices and get a hold on their smartphones again, they instantly change back to the normal anxious state and the core ideas of thinking calmly and thoroughly are abandoned. This phenomenon seems to argue that in order to keep ourselves mindful, we need to get rid of the modern technologies. However, it is impossible to entirely keep technologies out of our lives. Instead, what we should focus on is to find out the correct balance of using technologies that ensure the mindfulness of ourselves.

The University of Virginia is trying to build a contemplative collaboration space, which serves the purpose of flourishing students and improve the overall well-being of students. One main idea of the space is to provide students with a well-facilitated space that ensures good team dynamics and collaborative activities with promoting the idea of contemplation at the same time. My research question is to study what role different types of technologies can play in designing the space to achieve the ultimate goal of fulfilling students' needs and understand both the pros and cons in integrating technologies in our lives. In the literature reviews, I have researched about the relationships between three different pairs: contemplation and collaboration, technologies and contemplations, and technologies and room design. The discussion of technology use in the contemplative collaboration space will be consisted of a passive perspective, which is how personal devices influence us based on our values, and an active perspective, which is how to implement technology actively into our lives.

Literature reviews

Primarily, it is important to study the role contemplation plays when performing collaboration as we are trying to promote both at the UVA contemplative collaboration space. Contemplation refers to the action of profoundly thinking about something for a long time, which can be seen as an individual activity. While collaboration is an action of working together with someone or a group to achieve a goal. However, these two actions do align together. A study has shown that a good collaboration space also requires sufficient sense of personal privacy and concentrated minds (Congdon, Flynn, & Redman, 2014). At the same time, Sabel's study has found that with an appropriate set of contemplative practice, students will more willingly to perform critical thinking and better engage in classroom (Sable, 2014). Therefore, it is important to conserve the idea of solitude at the contemplative collaboration space.

Additionally, we need to learn about how technologies and contemplation can work together to simulate mindfulness. It is intuitive to understand that the presence of technology is irreplaceable in college group work activities. There are multiple interfaces that support the process of collaborating on a certain project. For example, message apps allow group members to send related resources to each others and everyone is able to easily access the contents on their laptops or smartphones, which gives them the capability to read at their own pace and revisit unclear sections. Another example is the collaborative online workspace like Google Docs/Sheet/Slides. With the help of these workspaces, everyone can work on the same file

simultaneously and contribute to the final product. Therefore, it is not surprising that different technologies are essential at collaboration spaces. On the other hand, contemplation and technologies are not put together normally. People imagine contemplative activities such and meditation is an action that is done without A study done by Nicky Duenkel presented the idea that when combining technology and contemplative together intentionally, students have improved skills in learning and thinking more thoroughly (Duenkel, 2013). Furthermore, with contemplative practice serving as a tool for calming minds and encouraging consciousness, students are able to "bring back" themselves from the inattentive mindsets caused by technologies. Similarly, Hirshberg argues in his article that contemplative pedagogy helps students return to a more focused and well-constructed mindset and thus become cognizant digital citizens (Hirshberg, 2019). Some interesting devices have also been developed to facilitate meditation practices, such as the device called Muse. According to its description, the device follows neuroscience concepts and through the advanced preprocessing signals, Muse interprets the user's mental activities to guide the user's contemplation practices ("Neurofeedback EEG Device - How it Works"). Contemplative activities and technologies will work best when they are designed in the correct way. As we are designing possible components of the contemplative collaboration space, it is essential to give attention of integrating the two ideas together and making students conscious about balancing well between the digital world of haste and the focused calmness of minds.

Since the group project is focusing on a place design process, I am also investigating on the space designing parts that can be related to technology other than personal technological devices, such as lighting and temperature control systems. A study conducted on Dutch elementary pupils has indicated that with the correct setting of brightness and CCT for the lights, the students perform better than students in traditional setting of lights (Sleegers, et al., 2012). Moreover, studies have found that uncomfortable room temperature distract people from their work and decreasing the temperature of an office from a too warm setting improves the average productivity (Clements-Croome, 2018). Therefore, paying close attention to what technological advances can be appended will help give students an overall better environment that allows contemplation and collaboration happen simultaneously.

STS Framework

In order to understand how technology can facilitate the contemplative collaboration space and analyze in a STS framework, I will use the ideas from the STS research done by Wetmore on the Amish Community, which examines the values of technologies, and the mindful use of technologies in the building of Amish Community from an STS analysis perspective (Wetmore, 2007). They presented extreme cautions when dealing with technologies and believed technology can reinforce social norms. There are two main principles in using technology in Amish society. Firstly, the technology chosen must be able to promote the Amish values instead of weakening them. Additionally, the Amish try to make sure that they are using technologies that are different than the rest to keep their uniqueness. This idea is an inspiration for thinking about the role of technology at the building, which means we need to pay attention to decide what we use and how we use. It is important to understand what exactly the values of technology are for UVA students and faculties, as well as how we will use the values to position the role of technology in our UVA community.

Methods

In order to get a better sense of how people value technology and what they would like to use, I plan to include data from different ways: short questionnaires and detailed interviews in person. Firstly, I need to conduct a questionnaire with students and faculties. In the questionnaires, I will ask what their opinions on technology, their habits of using technology, and how technology has influenced their lives. Some example questions can be the number of times they unlock their screens during normal learning/working hours, which kinds of technological devices they use more often, and what places they use technological devices frequently. Learning the students' relationship with technology will help inform me about how we should treat electronic devices in the contemplative space area from the passive perspective. For the detailed interviews that are conducted by face-to-face interaction, I will ask them to talk about what types of benefits/shortcomings have been brought by their use of technology. Moreover, I will let them write about what the space should have in it, specifically focusing on technological components, which is the active perspective of implementing technology in the new space. Moreover, I plan to gain some professional help from interviewing Professor Ehsan Baharlou from the School of Architecture to ask about the architectural view on technological implementation in room designs and how that influences the user experience.

Discussion and Next Steps

Combining the information learnt from students' and faculties' needs and opinions from questionnaire and interviews with the ideas from the professional views in architectural perspective, I will be able to understand people's use of technology is active and passive ways. The combined and comprehensive idea will guide me to bring in components and ideas I found from literature reviews so that I can successfully produce a report that understand the current roles of technologies and how we value them and make proper use of them to promote mindfulness and collaboration by both active and passive adjustments.

Bibliography

Clements-Croome, D. (2018). Creating the productive workplace: places to work creatively (2nd ed.).

Congdon, C., Flynn, D., & Redman, M. (2014, October). Balancing "We" and "Me": The Best Collaborative Spaces Also Support Solitude. *Harvard Business Review*. Retrieved from https://hbr.org/2014/10/balancing-we-and-me-the-best-collaborative-spaces-also-support-solitude

Duenkel, Nicky (2013) "Balancing Inner and Outer Technologies: Mobile Devices and Contemplative Practice," Journal of Pedagogy, Pluralism, and Practice: Vol. 6 : Iss. 1 , Article 4. Retrieved from: https://digitalcommons.lesley.edu/jppp/vol6/iss1/4

Hirshberg, D. A. (2019, June 18). Contemplating the Smartphone Dis/Connect. Retrieved November 6, 2019, from http://rsn.aarweb.org/spotlight-on/teaching/contemplative-pedagogy/smartphone-disconnect.

Neurofeedback EEG Device - How it Works. (n.d.). Retrieved from https://choosemuse.com/how-it-works/.

Sable, D. (2014). Reason in the Service of the Heart: The Impacts of Contemplative Practices on Critical Thinking. *The Journal Of Contemplative Inquiry, 1*(1). Retrieved from https://journal.contemplativeinquiry.org/index.php/joci/article/view/2/12

Sleegers, P., Moolenaar, N., Galetzka, M., Pruyn, A., Sarroukh, B., & Zande, B. V. D. (2012). Lighting affects students' concentration positively: Findings from three Dutch studies. *Lighting Research & Technology*, *45*(2), 159–175. doi: 10.1177/1477153512446099

Wetmore, J. (2007). Amish Technology: Reinforcing values and building community. *IEEE Technology and Society Magazine*, *26*(2), 10–21. doi: 10.1109/mtas.2007.371278