Thesis Portfolio

Fall Risk Classification Among Seniors

(Technical Report)

Big Data and Privacy: Finding the Balance in Distrust and Progress

(STS Research Paper)

An Undergraduate Thesis

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

> > Ziyang Guo Spring, 2020

Department of Biomedical Engineering

Table of Contents

Sociotechnical Synthesis

Technical Report Title

Big Data and Privacy: Finding the Balance in Distrust and Progress

Thesis Prospectus

Sociotechnical Synthesis

The following portfolio contains two related projects. The first project is technical research that focuses on utilizing patients' gesture data to predict the likelihood of falling in the subsequent 12 months. To ensure the success of the research, my team was given data, including patients' private information such as age, sex, and medical history. The excess data from this project led me to realize that researchers all over the world are most likely receiving information that infringes personal privacy of the subjects studied. Therefore, I decided that I wanted to learn more about privacy in research in the STS research paper portion of this portfolio. During my time digging into data privacy complication in research and in the past, I found out that our personal information has become a commodity that is frequently traded on the internet without our knowledge and permission. The degree of privacy infringement and the disregard of decency in selling people's private information caused me to divert my research focus to determining the underlying reasons for so many companies and entities to commit such a dishonest act. Therefore, I eventually landed my STS research on history case studies that illustrate regulatory failures in controlling personal privacy and the lack of public knowledge in the undergoing treachery.

My capstone technical project consists of developing a novel approach in predicting the likelihood of a patient falling in the subsequent 12 months after testing. Fall risk analysis has always been a challenge in the medical field due to the underdevelopment in the related technology. Many scholars and researchers have used statistical or stochastic methods to tackle fall risk prediction with limited success. My team proposed to use neural networks, which are technology that is often used in machine learning and artificial intelligence, to carry out the prediction. Two neural networks are introduced and implemented in my project, long short-term

memory (LSTM) and convolutional neural network (CNN). LSTM is implemented to understand the temporal relations between the sequence of gait features and the risk of falling. CNN is implemented to seek insight into the spatial relations of the gait features at each stage to the risk of falling. These two algorithms are designed to allow computers to understand the correlations of a patient's gait to his likelihood of falling. This project could be used as prototype reference in the future development of risk analysis and probability prediction in the medical field.

My STS research focuses on two historical data breach cases. The first case involves analysis of Anthem data breach, which is the biggest data leak in medical insurance. In the demonstration of development of the Anthem case and its consequent results, I want to show that the lack of regulatory response in the required defensive measure of cyber-attacks is one of the most important root causes in the current rampant data breach. The fast progression of informational technology mandates companies and entities to take actions to protect consumers' data privacy. Yet, regulators are severely lagged behind to implement a system or an agenda that upholds the standard of defensive measures and handles the potential repercussions. The second case is the story of graphic leakage of Facebook accounts in which the company twisted the understanding and meaning of well-implemented regulations to cheat the system. In this second case, I want to show that the most advanced regulatory system still has flaws and is in need of improvement. Most importantly, I want to compare the differences between the U.S. regulations and that of Europe to present the gap in data safety of two similar worlds. Throughout this project, I want to broaden the knowledge of what is happening to our data safety and what is lacking in protecting it.

My technical research project is a catalyst of the conception of my STS research project. My technical research gave me basic understanding in the field of medical research and its potential dilemmas, which fuels me with inspiration of going deep into the issues of data privacy. If these two projects were taken independently, I would not have considered data privacy as an important aspect of sociotechnical problems that was urgent to be addressed. My STS project expands my sight of one of the critical problems in the real world and continues to inspire me to delve deeper into issues that have skipped my mind. My STS project also aids with my technical project by providing a more thoughtful and conscious mind in doing the research. My projects made me become more mindful of other's privacy and the problem that related to privacy infringement in the recent years.