Temperature Checkpoint System
(Technical Report)

Mitigating Harmful Machine Learning Dependencies
(STS Research Paper)

A Thesis Prospectus Submitted to the

Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements of the Degree
Bachelor of Science, School of Engineering

Gregory Victor Vavoso

Spring, 2017

Technical Project Team Members
Matthew Bain
Andy Hui
Amanda Rein
Jack Schefer
Gregory Vavoso

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Signature ____ Date __________
Gregory Victor Vavoso

Approved ____ Date __________
Harry C. Powell, Department of Electrical and Computer Engineering

Approved ____ Date __________
Sean Ferguson, Department of Engineering and Society

Commented [SF1]: put in appropriate department/school affiliation

Introduction

The temperature-checkpoint door network we are designing will collect the temperature of a person who has approached the door (without physical contact), then unlock the door to allow the person's entry if their temperature is inside of the normal range (i.e. not feverish). The system will also publish these temperature readings to an online dashboard, where interested parties can see trends in the number of people approaching the door and their temperature readings. As such, the system will prevent people with fevers from entering the space, whether it be an office, classroom, or retail establishment, which protects the health of those within.

The sociotechnical research portion of this project addresses humanity's increasing dependence on machine learning systems. This research aims to develop a fundamental set of useful precautions and harmful dependencies that humans must abide by when determining their level of interaction with machine learning systems. This research will take an actor-network theory approach to understanding various crowdsensing human machine networks. Understanding the nature of the interaction between human actors and machine actors will assist in understanding the harm in becoming overly dependent on machine learning systems.

This temperature-checkpoint door network (technical project) is an example of a basic machine actor that could be investigated by the STS portion of research. The interaction between humans and this temperature monitoring system present an example of a human-machine actor-network that is highly coupled with pandemic behavior. The data analytics dashboard that is being created in this technical project can be used to better understand human behavior in this network.

Technical Topic

The goal of this technical project is to help combat the spread of the novel coronavirus, also known as COVID-19, as well as future illnesses. As economies start to reopen, concern for public health has given rise to the demand for technology to find and monitor cases (Morrissey, 2020). One example technology is the various contract tracing mobile apps that have emerged (COVIDWISE, 2020). These apps keep track of nearby devices in order to build up a network of which users have been in close proximity to other users. Thus, when a user tests positive, they can privately and anonymously notify all other devices that were nearby, i.e. the people that they may have been in contact with. However, these technologies are limited by the number of users they can attract and the willingness of these users to report positive results. One of the major goals of this project is to create a technology that does not rely on user motivation. Rather than having users actively enroll in public safety measures, the default state should be participation.

The logical next question is who would have the incentive to ensure public safety, as well as the power to enforce safety measures? Brick-and-mortar businesses like restaurants and retail stores are likely candidates. Their motivations are two-fold. First, in order for customers to return, they must feel safe, so increasing safety steps may drive in-person traffic. Second, safety measures protect themselves and their employees from the virus and from potential legal liability (Elejalde-Ruiz, 2020). Additionally, since these are privately owned businesses, they can generally enforce limits on who can enter their property. Other possible groups that may adopt this particular system are health care centers and educational institutions. Health care centers such as the UVa Elson Student Health Center have been asking visitors to wait outside the doors until someone can come manually take their temperature before admitting them. This relies upon visitors having the patience to wait outside, rather than entering through one of the entrances. It is also an inefficient system, especially for the staff of the health care center who have to constantly staff the entrance. As for why an educational institution would be interested in this system, study spaces or rooms at the fitness centers would greatly benefit from a system installed on the doorways. Therefore, we are designing an automated temperature screening system to admit people to the sensitive spaces described above. This system would be attached to the door-frame or as a separate kiosk, and temperature scans would occur without requiring physical device contact. If the temperature is within a predetermined bound, the user will be admitted to the facility. That way, they can similarly cut back on the human interaction going into having a person measure your temperature. This device will synchronize with a web application that tracks usage and temperature rates for the owners of the space to have a holistic view of the door usage and allow cleaning measures to be focused appropriately. Also, the web application's publication of the number of people that have been admitted to the space would be particularly helpful for students who may want to avoid peak times at the gyms or libraries.

Item	**Total Price**	**Product Page**
Raspberry Pi 3B+	$45.00	Digikey
HC-05 Bluetooth module (x2)	$12.95	Mouser
IR Temperature Sensor	$31.36	Digikey

Fail Safe Lock	$19.85	Walmart
32 GB SD Card	$13.49	Amazon
PCB Estimates	$60.00	N/A
Housing material	$50	N/A
Replacements/Misc.	$100	N/A

Table 1. Resource breakdown

The success of this project is dependent on meeting our deliverable requirements. Our project has 3 main deliverables: a temperature sensing system that accurately records temperatures, a door lock that opens on command, and a web dashboard that compiles the information retrieved by the door lock and the temperature sensor. These three components must fit within a $500 allotted budget. The resource requirement breakdown can be seen in Table 1.

My personal responsibilities include overall device integration and assembly as well as embedded code design. Overall device integration and assembly entails managing each subsystem and defining requirements necessary for cohesive subsystem interaction. Embedded design entails programming a real-time operating system (RTOS) to produce desired functionality (FreeRTOS, 2020).

STS Topic

Machine learning systems are increasingly becoming a part of everyday human life. Each year, we become more dependent on machine learning models to advise us. For example, many individuals are highly dependent on navigation apps to get from point A to point B. As of 2018, nearly 77% of smartphone users regularly depend on navigation apps to travel (Panko, 2018). This drastic increase in dependence may be cause for concern (Anderson & Rainie, 2018).

Evidence has shown that many software applications have been affected by this spontaneous change in human behavior. Specifically, many artificial intelligence and machine learning applications have been unable to handle this sudden change in human behavior. For example, Amazon's previous top searches were replaced with COVID-19 related products such as face masks, hand soap, and cleaning wipes, causing previous models to perform in unexpected ways (Heaven, 2020). Machine learning models are designed to react to small changes in human behavior. However, most models are only trained on past data. Models begin to fall apart when global human behavior takes a sudden turn. While Amazon shopping suggestions may not seem like a dire cause for concern, there are machine learning models that are life critical, such as traffic models for self-driving cars. For example, the Google Maps team had to completely overhaul their traffic models, as they were unable to rely on old data that did not represent the new pandemic behavior (Quach, 2020).

In order to conduct this analysis, an actor-network theory approach will be taken to analyze the human-machine networks (HMNs) involved in this pandemic (Tsvetkova, Understanding Human-Machine Networks: A Cross-Disciplinary, 2017). The two actors in consideration are humans and machines. The human users will be the primary stakeholders, specifically the humans receiving critical information from machine learning systems. For simplicity of analysis, humans can be an individual person or groups of individuals. In most cases, machines will refer to a single machine learning system, such as Google Maps.

Human to machine interaction will be classified as either passive contribution or active contribution. For example, Google Maps users providing location data to Google would be a passive contribution. Machine to human contributions will mainly be defined as active contributions, such as Google Maps giving a user specific direction.

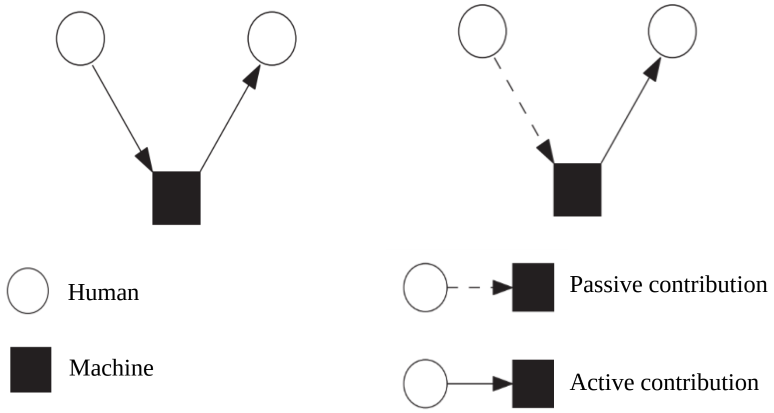

Figure 1. Crowdsensing human-machine network

Thus, the specific analysis will be centered around “crowdsensing” human-machine networks (Tsvetkova, Understanding Human-Machine Networks: A Cross-Disciplinary, 2017). This crowdsensing form of HMN is shown in Figure 1. Human contribution in a crowdsensing network can be either passive or active, as shown by the two networks in Figure 1. The left network shows active contribution from humans, while the right network shows passive contribution from humans. Machine to human contributions tend to be active contributions in crowdsensing networks.

In order to ground analysis in a real-world and life critical application, traffic patterns and self-driving artificial will be used as the primary crowdsensing network in consideration (Kellner, 2019). However, other applications will be mentioned to show the broad scope of this issue.

Depending on the path this research takes, the overarching objectives of this analysis may be twofold. A first objective would be to enumerate a set of fundamental precautions humans must take when designing and relying on machine learning models. Secondly, it may be discovered that certain dependencies on machine learning systems must be avoided at all costs and therefore a list of dependencies to avoid could also be stated.

The investigation will aim to justify and create a basis for understanding these fundamental precautions and dependencies. An idealistic goal would be to outline two to three fundamental principles that software engineers and regular users should abide by and consider when interacting with machine learning systems in order to prevent catastrophic failure and loss of life. However, a more realistic goal would be to simply identify and discuss potential harmful dependencies and useful precautions.

Evidence gathering will consist of online research, interviewing machine learning subject matter experts, performing first-hand data analysis, and gather any other sources that can provide insight into the severity of this issue. I plan to use data analytics techniques myself to gather pre-pandemic and post-pandemic data to create my own conclusions on how drastic the changes in human behavior were. Additionally, gauging the concern of professors and software engineers who are experts in the field of machine learning will be important in creating a set of precautions and harmful dependencies. Interviewing and contacting subject matter experts will be a great way to find further research entry points into this analysis.

There are some key questions that will guide this research process. Some examples include: Is complete catastrophe due to machine learning dependence inevitable? What levels of catastrophe exist? Is it possible to create a finite set of principles to completely avoid catastrophe? What are the limits to how much humans can depend on machine learning systems? While many of these questions may not be answerable in the scope of this research, they will act as guiding questions for developing a set of useful precautions and harmful dependencies.

Next Steps

- Near future
  - (Technical) Complete subsystem testing and assemble unit
  - (Technical) Undergo extensive assembled product testing to ensure proper temperature accuracy
  - (STS) Begin contacting interviewees for further investigation
  - (STS) Create a plan and structure for data analysis of pandemic behavior
- Distant future
  - (Technical) Complete final write-up of technical project
  - (STS) Carry out interviews and deeper research into topic
  - (STS) Finalize precise expectation for how many precautions and dependencies to identify
  - (STS) Gather evidence into cohesive groups
  - (STS) Enumerate and refine final fundamental precautions and harmful dependencies to avoid

References

Anderson, J., & Rainie, L. (2018). Artificial Intelligence and the Future of Humans. *Pew Research Center*.

Automated Vehicles for Safety. (2020). Washington, DC, United States: National Highway Traffic Safety Administration.

Bathaee, Y. (2018, Spring). THE ARTIFICIAL INTELLIGENCE BLACK BOX AND THE. Cambridge, Massachusetts, United States.

Campbell, E., Sittig, D., Guappone, K., Dykstra, R., & Ash, J. (2007). Overdependence on Technology: An Unintended Adverse Consequence of Computerized Provider Order Entry. National Center for Biotechnology Information.

COVIDWISE. (2020). Retrieved from Virginia Department of Health: https://www.vdh.virginia.gov/covidwise/

Elejalde-Ruiz, A. (2020, May 4). If you get sick with COVID-19, is your employer liable? As businesses prepare to reopen, worker safety is a priority. Chicago, Illinois, United States.

Fisher, M. (2013, April 23). Syrian hackers claim AP hack that tipped stock market by $136 billion. Is it terrorism? Washington D.C., United States: The Washington Post.

FreeRTOS. (2020). Retrieved from FreeRTOS: https://www.freertos.org/

Hancock, P., & Nourbakhsh, I. S. (2018, April 16). On the future of transportation in an era of automated and autonomous vehicles. Pittsburg, Pennsylvania, United States: Proceedings of the National Academy of Sciences of the United States of America.

Heaven, W. D. (2020). Our weird behavior during the pandemic is messing with AI models. *MIT Technology Review*.

Hosanagar, K., & Cronk, I. (2018, October). Why We Don't Trust Driverless Cars - Even When We Should. Cambridge, Massachusetts: Harvard Business Review.

Kellner, L. (2019). Machine Learning Algorithms Help Predict Traffic Headaches. *Berkely Lab*.

Long, B. (2020). The Ethics of Deep Learning AI and the Epistemic Opacity Dilemma. APA Online.

Maguire, E., Gadian, D., Johnsrude, I., Good, C., Ashburner, J., Frackwiak, R., & Frith, C. (2000). Navigation-related structural change in the hippocampi of taxi drivers. Montreal, Canada: Proceedings of the National Academy of Sciences of the United States of America.

Morrissey, J. (2020, June 16). Fighting the Coronavirus With Innovative Tech. *The New York Times*.

Panko, R. (2018). The Popularity of Google Maps: Trends in Navigation Apps in 2018. Wasington, D.C., United States.

Quach, K. (2020). Google declares Maps COVID-19-ready after retraining it on pandemic traffic – or the lack of it in some areas. *The Register*. Retrieved from https://www.theregister.com/2020/09/03/google_maps_covid/

Rudin, C., & Radin, J. (2019). Why Are We Using Black Box Models in AI When We Don’t Need To? A Lesson From An Explainable AI Competition. MIT Press.

Steinhardt, J., & Toner, H. (2020, June 8). Why Robustness is Key to Deploying AI. Brookings Institution.

Thomas, M., Norton, J., Jones, A., Hopper, A., & Ward, N. (2011, March). Global Navigation Space Systems: reliance and vulnerabilities. London, United Kingdom.

Tsvetkova, M. (2017). Understanding Human-Machine Networks: A Cross-Disciplinary. *ACM Journals, 50*.

Tsvetkova, M., Yasseri, T., Meyer, E. T., Pickering, J. B., Engen, V., Walland, P., . . . Bravos, G. (2017). Understanding Human-Machine Networks: A Cross-Disciplinary. *ACM Journals, 50*.

Verbeek, P.-P. (2006, May). Materializing Morality: Design Ethics and Technological Mediation. Enschede, Netherlands.