

Digital Storage of Personal Identification Documents
(Technical Paper)

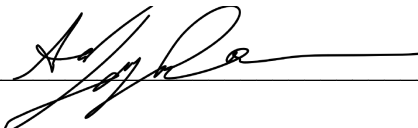
The Application of Heuristic-Based Neural Networks and the Dangers of Bias
(STS Paper)

A Thesis Prospectus
In STS 4500
Presented to
The Faculty of the School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science


By
Sridhiraj Jayakumar
November 1, 2019

Technical Project Team Members:
Eric Burbach
Christopher Han
Samantha Kostelni
Gio Lee
Amanda Murray

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:  _____

Approved  _____ Date 11/15/20
Michael Gorman, Ph.D., STS Program Director, Department of Engineering and Society

Approved  _____ Date 11/25/19
Ahmed Ibrahim, Ph.D., Assistant Professor, Department of Computer Science

I. Introduction

Intelligence is defined as the ability to learn, understand, and apply knowledge towards a problem in order to develop a solution. It is generally applied in order to solve, reason, and learn different situations. Some of these situations involve integrating “various cognitive functions such as; language, attention, planning, memory, perception” (Shabbir & Anwar 2015). Typically, intelligence applies to living organisms that leverage their knowledge in order to discover solutions. However, in more recent years, this general “problem solving” intelligence has been extended to computers, and it has been coined artificial intelligence or AI. Many industries have started to incorporate AI into their algorithms and its techniques, specifically heuristics, will have a significant impact on society.

For my technical topic, I am working with my team, with 6 other group members, in order to develop a digital storage application, Identityti, that will house documents that are generally used for personal identification. The long-term goal of the application is to eliminate the need for users to physically keep track of important documents, and to securely store and use them at established institutions. A simple use case would be to store the documents needed to apply for a driver’s license (i.e. social security card, proof of residency, etc.) and use them at the DMV.

One of the goals of our project is to use AI in the form of image recognition software in order to determine what kind of document the person has uploaded. This eliminates the need for the user to manually enter that information, and eliminates the need for the admins to manage this information.

For my STS topic, I will research the different settings in which AI is being used and the societal implications AI can have under the STS framework of ethics. Specifically, I will

investigate common pitfalls of heuristic algorithms, and determine the benefits and drawbacks of using these techniques.

II. Technical Topic

Identityti is a document storage service for documents that otherwise typically only exist as a hardcopy. The client shared several user stories that illustrates the problem that the service is solving. The user stories highlight how users tend to lose physical documents, spend more time shuffling through physical documents, spend more time organizing physical documents, and cause more hassle when attempting to authenticate physical documents.

Because the documents uploaded to this platform is typically highly sensitive, data security will be a critical challenge to tackle. Login authentication would be performed by Auth0 which already provides an array of mitigating web-app security breaches such as anomaly detection and force email verification (Poza 2018). Additionally, in order to allow clients to access their data from anywhere, the documents they upload will be stored in the cloud, specifically in an AWS S3 bucket. When uploading the documents, the application will save certain metadata associated with the document allowing the client to easily search/sort/categorize the document, making it convenient and faster to find it later. The International Data Corporation, who is a provider of market intelligence, conducted a study on their workers to gauge how much time they spend weekly looking for physical documents. In a group of 1200 workers, IDC found that “they spend an average of 4.5 hours a week looking for documents” (Biddle 2017). Since Identityti is targeted at both consumers and enterprise clients alike, searching for these documents on Identityti, whether the client is an individual or a business, will be much faster than searching for physical documents.

One concern for storing sensitive data in the cloud is that these services can be compromised. However, in order to mitigate this issue, the plan is to encrypt the data stored in S3, so that even if the bucket is compromised, only encrypted data can be recovered. By storing these encrypted files in the cloud, Identityti offers a secure and fast solution for clients, allowing them to easily share their documents with enterprises.

Identityti also provides benefits for enterprise clients. A major problem that large enterprises face is the sheer size of information they have to process and handle. Using Identityti, enterprises have the ability to create an account in order to manage all of their employees' data. For example, when onboarding a new employee, an enterprise client could request the necessary documents, such as identification and tax reports, from the new employee through Identityti. In this way, Identityti creates a secure and simple path of communication between enterprises and employees for personal, confidential documents. Instead of requiring users to carry physical documents and submit them to enterprises, Identityti creates a channel to share these documents electronically.

Storing physical documents is stressful and confusing. It is seemingly ambiguous as to what one should keep and what one should throw away, and safeguarding the ones that are kept requires a high level of organization. Identityti alleviates this problem for its users by recommending to the user which documents they may want to obtain and store on the service. Depending on what documents a user has already stored, Identityti will provide feedback on what important documents are missing, as well as what documents are required for common tasks. For example, when attempting to obtain a driver's license, Identityti will tell the user exactly what other documents are needed in order to apply for a driver's license.

III. STS Topic

A. History

Artificial intelligence is the ability of computer systems to perform tasks, that normally require human intelligence. This may include visual perception, speech recognition, decision-making, etc.

One of the first projects, which is regarded as the birth of AI, was performed in 1942 by the English mathematician Alan Turing. He developed the code breaking machine (The Bombe), used during World War II, that helped decrypt the Enigma Code used by the Germans (Haenlein & Kaplan 2019). This code had roughly $1.5 * 10^{19}$ unique combinations, which made it nearly impossible for humans to decode physically, especially since the Germans changed the code daily. However, with the help of Turing's *Bombe*, they were able to crack the code and, essentially, help win the war (Saranji & Sharma 2018). The machine worked by following a similar algorithm used by the Enigma Code to encrypt messages, and ran through the different possibilities (Wilcox 2006). While, the Bombe simply used complex algorithms and the brute force method in order to break the code, this was the start of developing more complex AI machines.

Another significant milestone in the field of AI came from IBM's Deep Blue chess playing program in 1997. This machine managed to win against the world champion Gary Kasparov, showing that these machines could be used to optimize different processes (Wilcox 2006). However, Deep Blue was also a "simple" machine, in that it did not have the capacity to learn and simply calculated all possible moves beforehand to choose the best move.

It wasn't until 2015, that significant progress had been made in the AI field. This progress came in the form of a machine called AlphaGo. AlphaGo is a program that played the game of Go, which is a strategy board game, similar to chess, invented in China more than 2500 years

ago. However, Go is hundreds of times more complex than chess, in terms of the number of possible moves a player can make. For example, after the first two moves of a chess game, there are 400 possible next moves, while in Go, there are 130,000 (Muio 2016). Even with this staggering number of possible moves, AlphaGo managed to beat the world champion, Ke Jie, in 2017. This achievement was important step in developing better AI because AlphaGo utilized a process called artificial neural networks. These networks mimic the process of neurons in the brain, and allow computers to heuristically gain knowledge from test data and previous game data. This essentially means the computer learns as it plays. Computers no longer needed to know all the information about a problem; And instead of using brute force, they could develop techniques themselves in order to solve a problem. With this substantial breakthrough, many doors opened for the future of AI.

B. Learned Bias

Currently, AI is being used in many ways, especially to optimize processes in society through the use of heuristic neural networks. Some examples of where this technology is being used include self-driving cars, HR hiring processes, stock market predictions, etc. With the integration of this novel technology, everyone, from the public to business owners, will be impacted. By viewing this topic through the STS framework of ethics, and exploring how normalized deviance can cause ethical concerns, it might offer some insight regarding the benefits and disadvantages of using neural networks.

To deal with the large influx of data companies receive, they turn to artificial intelligence and machine learning in order to review this data. They can be used to review information ranging from financial data, such as evaluating credit for loan applications, to personal data, such as determining if an applicant is the right fit for a company (DeBrusk 2018). Cloud platforms even

offer a “machine learning in a box” service, that allow users to build their own machine-learning models. Essentially, these models use neural networks and train against test data to determine parameters that can be used to classify data. With this high accessibility and cost-effective method, many people can take advantage of a machine’s computing power in order to make models that suit their needs. However, these models are only as good as the data they are trained against.

These intelligence networks are susceptible to something called the “garbage in, garbage out” syndrome. For these neural networks, “the type of ‘garbage’ is biased data. Left unchecked, feeding biased data to self-learning systems can lead to unintended and sometimes dangerous outcomes” (DeBrusk 2018).

For example, in 2016, Microsoft developed a chatbot named Tay, modeled to sound like an average teenage girl, that would respond to people on Twitter. According to Microsoft, the bot would learn from those who communicate with it. The experiment started out fine, with Tay tweeting messages like “humans are super cool” and “c u soon humans need sleep now so many conversations today thx”. However, it soon started to devolve as people started tweeting racial slurs, offensive words, remarks, and ideologies. Tay learned from this public data and started tweeting similar tweets (Coval 2018). Soon after, Microsoft shut down Tay to avoid it from tweeting more offensive messages. This experiment highlights how AI can be manipulated into performing unintended actions. With the influence of this “garbage” data, Tay learned how to tweet like its audience, resulting in hate speech and belligerent behavior.

Another example of bias in artificial intelligence comes from a program developed by the company Northpointe. This company developed the Correctional Offender Management Profiling for Alternative Sanctions, or COMPAS. This product is used in some courts to perform

risk analysis on defendants and is based upon major theories of criminality, including criminal personality, social isolation, substance abuse, and residence/stability. Using these categories, defendants can be ranked low, medium, or high risk for committing a repetitive offense. Many jurisdictions adopted this sort of risk analysis, as an additional measure used to determine a ruling against a defendant. However, statistical analysis conducted by ProPublica show that COMPAS may be racially biased when determining risk scores. ProPublica found that the “tool correctly predicts recidivism 61 percent of the time. But blacks are almost twice as likely as whites to be labeled a higher risk but not actually re-offend. It makes the opposite mistake among whites: They are much more likely than blacks to be labeled lower risk but go on to commit other crimes” (Angwin & Larson 2016). In one case, defendant Zilly scored as high risk for violent recidivism and was sent to prison with a 2-year sentence. Later, a public defender appealed this sentence and called upon the score’s creator, Brennan, as a witness. After explaining that COMPAS shouldn’t be used as sole evidence in a ruling decision, the judge reduced Zilly’s sentence to 1 year and 6 months. Later the judge commented, “Had I not had the COMPAS, I believe it would likely be that I would have given one year, six months,” showing how much influence COMPAS had on the judge’s original decision (Angwin & Larson 2016). Given incomplete data to train on, COMPAS had identified race as an input parameter, giving it an inherent racial bias. Using AI in court systems has a potential to seriously alter the course of many people’s lives either by falsely classifying defendants as high risk or influencing the courts’ decision, and, as such, should be used cautiously.

IV. Conclusion

Artificial intelligence is a powerful tool in modern society. It can be used to optimize many systems and help processes become more efficient. It has become so cost-effective, that we can

even use it in simple cases, such as classifying documents in Identity. However, when using these complex machines, society must be aware of the potential biases it can develop and strive to prevent it from learning from prejudiced data. By analyzing the process of how it can develop normalized deviance, artificial intelligence made in the future can try and avoid this problem and be less biased than its predecessors.

References

- Angwin, J., Larson, J., Kirchner, L., & Mattu, S. (2019, March 9). Machine Bias. Retrieved November 1, 2019, from <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
- Biddle, P. (2018, January 24). Productivity, Lost Time, and the Power of AI to Make Search Easier. Retrieved October 31, 2019, from https://medium.com/@diamond_io/productivity-lost-time-and-the-power-of-ai-to-make-search-easier-a59d4cd85a26.
- Coval, T. (2018). Artificial Intelligence, Bias & Your Business. *Journal of Property Management*, 83(2), 6–9.
- DeBrusk, C. (2018, March 26). The Risk of Machine-Learning Bias (and How to Prevent It). Retrieved November 1, 2019, from <https://sloanreview.mit.edu/article/the-risk-of-machine-learning-bias-and-how-to-prevent-it/>.
- Desmond, J. (2018, March 13). Normalization of Deviance Endangers AI Self-Driving Cars. Retrieved from <https://www.aitrends.com/ai-insider/normalization-of-deviance-endangers-ai-self-driving-cars/>.
- Haenlein, M., & Kaplan, A. (2019). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *California Management Review*, 61(4), 5–14. doi: 10.1177/0008125619864925

- Muioio, D. (2016, March 10). Why Go is so much harder for AI to beat than chess. Retrieved from <https://www.businessinsider.com/why-google-ai-game-go-is-harder-than-chess-2016-3>.
- Neapolitan, R. E., & Jiang, X. (2018). *Artificial intelligence: with an introduction to machine learning* (2nd ed.). Boca Raton (Fla.): CRC Press, Taylor & Francis Group.
- Osegi, Ndidi, E., Anireh, & Ike, V. (2017, January 3). Deviant Learning Algorithm: Learning Sparse Mismatch Representations through Time and Space. Retrieved November 1, 2019, from <https://arxiv.org/abs/1609.01459>.
- Poza, D. (2018, July 19). How Auth0 Makes Your Apps More Secure. Retrieved October 31, 2019, from <https://auth0.com/blog/how-auth0-makes-your-apps-more-secure/>.
- Sarnagi, S. & Pankaj S. (2019). *Artificial Intelligence: Evolution, Ethics and Public Policy*. London: Routledge.
- Şenyiğit, E., Düğenci, M., Aydin, M. E., & Zeydan, M. (2013). Heuristic-based neural networks for stochastic dynamic lot sizing problem. *Applied Soft Computing*, 13(3), 1332–1339. doi: 10.1016/j.asoc.2012.02.026
- Shabbir, J., & Anwer, T. (2015). Artificial Intelligence and its Role in Near Future.
- Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Driessche, G. V. D., ... Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*. doi: 10.1038/nature16961
- Wilcox, J. E. (2006). *Solving the enigma history of the cryptanalytic bombe*. Fort George G. Meade, MD: Center for Cryptologic History, National Security Agency.