Thesis Project Portfolio

Security For Databases and Throughout Computer Science Curriculum
(Technical Report)

An Analysis of Machine Learning and Artificial Intelligence on Climate Change Through Capitalism

(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

Franceska Padilla Coo
Spring, 2024
Department of Computer Science

Table of Contents

Executive Summary

With the impending climate crisis, machine learning and artificial intelligence's (AI) impact must be addressed, with one of those ways being within academia and in practice. Database structuring and security are one way to address the overarching issue of data center waste and usage, and this past fall, I joined Professor Nada Basit in continuing to develop a database learning tool that reinforces proper database structuring and security. I personally was interested in proper and efficient database structuring, and I undertook this research in order to ensure other people learn this material early on and incorporate it into their practice to be proactive instead of reactionary with their data-based projects. Meanwhile, I began writing a research paper about the negative environmental impact of current machine learning and artificial intelligence's practices within the capitalist system. This analysis and research highlights the lack of warning and caution almost all sectors of society take when implementing machine learning and AI despite the fact that its contribution to carbon emissions is significant. One of the reasons why machine learning and AI have continuously been harming the environment is because of data center expansion, and with data center expansion, several other problems occur including database waste heat and data center construction. However, many data engineers and computer engineers lack an education in environmentally friendly and sustainable database and data center construction and even in the very fact that their development harms the environment in the first place.

When creating databases, not only security must be considered, but also the physical storage of the data plays a huge role in security and general data formatting. The way we keep data can either eat up at resources or can be optimized to be compact and secure. As the demand is anything but insignificant, resource allocation, structure, and security should constantly be the priorities with database creation and construction. The more compact data is, the less throughput on the traveling path for faster collection and retrieval. Well-constructed data minimizes the amount of data transfers and data traffic while still providing a lot of information quickly. Database security also falls under the subject of database construction. Data insecurities can render the data insignificant, skewing outcomes. If the collected, skewed data becomes the basis for any studies, results of the study could mean nothing or could lead to the implementation of something non beneficial or possibly even harmful. One famous instance of this was Amazon's implementation of a hiring AI, which heavily favored male applicants due to the lack of female applicants in the hiring pool. Data security ensures that this data remains encrypted and used only by those authorized to do so. Data insecurity could also provide data to the hands of the wrong people. This capstone used a pre-existing teaching application and built on top of it, allowing students to scan code and familiarize themselves with the syntax needed to execute three different sql injection protection methods. Using an additional prepared statement exercise, a syntactical exercise using double quotes, and a regular expression (regex) exercise, this capstone project aimed at providing students with interactive introductory practice problems, stepping through the order of operations and syntax for a robust set of sql injection prevention tactics that they can continue using in more practical environments.

I work on enforcing proper data storage, ensuring that data is grouped in a way for easy access and readability. Although common and incredibly basic, attacks on databases occur frequently because of poor management. Still, the structure of the data is well protected using preemptive measures that prevent easy data access. Methods like this include verifying the users attempting to access data, checking for malicious access requests, and restricting the types of

data entered. Even more, I encourage database designers to incorporate as many database design guidelines and security measures as possible to set a standard in database creation. Without these large databases, machine learning and AI would not have enough knowledge to learn and predict from. These operations require mass amounts of data to foresee patterns and account for different anomalies and learn. To continue collecting data, we must continuously prove the safety and proper management of data already on hand. The new exercises added consisted of three new exercises: a new prepared statement exercise, an escaped input exercise, and a regex check exercise. The user has to fill in blanks, implemented by creating an html form. By providing an explanation and encouraging users to fill in the blanks of pre-written code, the student gets to parse the entire structure and see where things are connected. It requires familiarity with the structure of the security measures. The more security throughout the entire development process can allow as many checks as possible to prevent security threats. This capstone can be continuously developed to provide positive reinforcement and security explanations depending on the username input. This learning application should be further developed to also include a section about environmental care and accountability. A data developer should walk in with knowledge of how much energy and natural resources data-driven technology eats up and know the practices to reduce this boundless usage. To ensure that these skills are taught early and therefore incorporated and encouraged early, this capstone project uses a pre-existing application for students in introductory databases. Additionally, the interface is interactive and calls for user participation through a hands-on activity with explanations, found to be effective among computer science students.

My topic surrounds the environmental impact of machine learning and AI, but I wanted to explore why the impact of machine learning and AI on the environment is not more well-known and stressed in public discourse, especially due to the increasing amounts of calls for environmental action on climate change. However, green computing and other attempts at reducing carbon emissions in technological production pose a threat to many producers as its initial implementation could potentially harm the immediate economy. This is even more applicable to machine learning and AI as their introduction to society has caused a stir from the public, and many producers hope to capitalize off their success, implementing their own versions of machine learning algorithms and AI. Society will probably forgo the health of the environment for the sake of AI's powers. Society therefore wants to maximize their benefit rather than scrap the technology. Again, this further indicates a priority of the economy, then technology like AI, and lastly, the environment. This problem will be analyzed using utilitarian ethics. The reason why utilitarian ethics will be used is because, based on initial research on the topic, the value of AI depends on how much the public thinks it will help them and how useful the results of machine learning and AI can be to them.

The minimal coverage of green computing contributes to this general lack of information. As Konig, Wurster, and Siewert discuss, society has handed over a sense of autonomy to AI since its introduction. Society, therefore, holds AI in a high regard, a higher regard than the environment. Notions like these impact our future as the climate crisis worsens, leaving society to "again find themselves at an important crossroad, where decisions taken today and behavioral patterns taking hold now may shape developments and environmental impacts of the technology over the next decades" (Konig, et al., 2022, 1). Using the example of "waste heat recycling" in European cities, Velkova demonstrates how society readjusts to keep using data centers and AI. Even more, Velkova claims that this importance stems from the fact that society acknowledges the investment of a significant amount of time and money in these technologies, "determined by

the interested parties in the production process" (Velkova, 2016, 4). Society therefore wants to maximize their benefit rather than scrap the technology. Machine learning and AI fall under technologies that could improve with green computing. However, because this benefit would require time and the adjustment of already established procedures, society prefers the more immediate economic benefit of forgoing sustainable procedures and proceeding with generation of new machine learning and AI technology. Even when society attempts to account for the environmental impact of machine learning and AI technology, the driving force remains economic in that consumers want to get as much value as possible from investments, regardless of negative environmental impact.

This research demonstrates a lack of acknowledgement of environmental impact by machine learning and AI in order to prioritize and optimize profit. This is shown through the reckless abandon society has approached machine learning and AI with, the lack of incorporation of environmental concerns in machine learning and AI education, and the approaches people took to address the few environmental concerns they did attempt to rectify. There is a greater battle and many more conflicting values in the business and corporate world once education is in application, but addressing database construction within education through effective learning applications like what I worked on in my capstone, can instill proper values and practices early on that remain incorporated throughout a person's occupational journey.