Artificial Intelligence: Using Machine Learning to Identify Dangerous Military Systems

CS4991 Capstone Report, 2023

Alex Joon Kim
Computer Science
The University of Virginia
School of Engineering and Applied Science
Charlottesville, Virginia USA
Ajk5qwb@virginia.edu

ABSTRACT

MITRE, a McLean-based non-profit company focused on serving the public interest, decided to automate the process of identifying potential hazards within a technical military system's design in order to make the tasks of System Safety Engineers more efficient and improve the safety of system users. I provided significant contributions to this effort by leveraging Keras, TensorFlow, Python, and other Machine Learning (ML) technologies to develop a program called the Potential Hazard Identification via Artificial Intelligence capability. This PHIAI solution implemented a variety of Machine Learning/Deep Learning models to improve upon previously implemented algorithms designed for this task. Although implementation is yet to be complete, the PHIAI capability has proven to be successful, consistently achieving impressive accuracies through multiple tests and applications. In its current state, the capability is especially useful when examining military system models but also has potential uses for systems in other fields as well. Looking towards the future, the technology must be finished and deployed, tested to ensure effective operation, and improved to become applicable to systems in other fields and sectors.

1. INTRODUCTION

Technical systems throughout the modern world are entrusted with important responsibilities by people who misunderstand the amount of risk associated with handing decisions over to technology. From the smallest of firms to the largest, critical information and tasks are being assigned to technical systems in order to effectively streamline business operations. While this has produced a variety of benefits over time, it has led to a sudden rise in concerning issues as well.

One such issue is the misuse or breakdown of technical systems. System malfunctions and misuse could result in immense financial losses, exposure of important data, and even the jeopardization or loss of human lives. While it is preferable to hope that these consequences never occur, these incidents are heavily prevalent today.

This issue is especially apparent in industries such as the military. Throughout daily military operations, government organizations have attempted to utilize technology to make military systems safer. Unfortunately, the effectiveness of this overall effort remains questionable at best. As industries and fields across the world continue to grow, the question remains: How can technology be used effectively early on to prevent accidents and losses from potentially dangerous military systems?

2. RELATED WORKS

One significant source of inspiration for my work was Lawrence (2002) who described the beneficial impacts of safety technology within the construction industry. The implementation of advanced safety technologies in the day-to-day operations of businesses in the construction industry has shaped them to make better decisions and drive efficiency to achieve safer projects. While Lawrence's work focused specifically in the construction industry, the influence of safety technology is applicable to all fields that foster dangerous work environments. The appliance of safety technology and potential to achieve safer operations in important industries such as the military is crucial for safety and aligns with MITRE's mission to protect national interests.

Another significant source of influence for my work was Nzongo (2023). Nzongo posited that there exists a relationship of mutual cooperation between humans and computers where each partner can benefit from the power of the other and thrive in a productive and viable partnership. The development of my project revolved around this concept of Man-Computer Symbiosis by ensuring that both the computer and human dimensions would exercise control over each other and perform tasks for which they were best suited. For the computer program, the PHIAI capability is entrusted with the responsibility of handling important data and correctly identifying potentially hazardous systems with efficient performance. The human users of the program are entrusted to supply the capability with productive data and ethically use the results of the capability for the benefit and safety of system users.

3. PROJECT DESIGN

The Potential Hazard Identification via AI (PHIAI) capability is comprised of four software components that work in conjunction with each other during execution. The first component and the base of the algorithm is a Python file named app.py. App.py contains all the functionalities involved within the algorithm including algorithm start-up, training, and hazard prediction. It also contains the Python packages and machine learning models used to create and run the algorithm. The second component of the capability is the machine learning model classes used within the app.py file. The third component of the capability is the Neo4j database querying/manipulation files used to extract and use military system data supplied by project sponsors. The fourth and final component of the capability are packages imported by app.py to improve the algorithm's performance, accuracy, or usability.

3.1 Requirements

The PHIAI capability is sponsored by officers within the military industry and is oriented towards System Safety Engineer clients who are charged with maintaining safety in complex and technical military systems. These clients required the capability to satisfy three important needs: accuracy, reliability, and understandability. The clients needed the PHIAI capability to indicate potentially dangerous subsystems accurately. Inaccurate results misrepresent the safety of inputted technical systems, resulting in financial and human consequences if these systems were to be misused or broken by unaware users. The clients also needed the capability's algorithm to maintain a consistently high performance (or be reliable) during its use. An unreliable algorithm will make it hard for users to entrust the machine to make correct decisions and in turn cause the capability to lose real life applicability. Finally, the clients needed the capability to be easily understood by its users. Otherwise, it would hinder the relationship of trust between the capability and its users and jeopardize the long-term applicability of the capability in safety preservation.

Despite the requirements, the PHIAI capability contains inherent system limitations that hinder its productivity. One such limitation lies in its lengthy machine learning model training times when being fit with training data. To ensure accurate results and productive use, the PHIAI capability must be trained extensively with a high volume of system/hazard data prior to its daily use by engineers, hindering its performance. Another system limitation lies in the errors that naturally arise in identification which lead to the mislabeling of certain subsystems. Due to the current limitations of predictive AI, it is very unusual for the technology to predict with flawless accuracy when being applied to a large set of unfamiliar data.

3.2 Key Components

The specifications of the project existed and was developed entirely within a software domain and through the Python programming language. The machine learning models were developed using Keras and Tensorflow, which are libraries dedicated to the development of AI in Python. To create and utilize endpoints, Flask was used as a micro web framework. The entirety of system training data was contained inside the Neo4j graph database. The input data used to train and test the PHIAI capability was relatively large in scale and featured a wide range of military technical system design data. Although additional GPUs were available for use in training, none were/are being used in its current state.

3.3 Challenges

One challenge in the development of the PHIAI capability was the initial lack of training data for the algorithm. Due to the limited amounts of training data supplied by the project's sponsors, training for the algorithm's machine learning models resulted in simple models that were fit too closely to the data. If the capability was released in this state, the algorithm would suffer when evaluating unfamiliar input data resulting in low identification accuracy and poor performance. Another challenge was that the model consistently achieved low accuracies when identifying hazards regardless of the lack of training data. This threatened the capability's applicability to real world data and violated the client needs for the capability.

Solutions to these challenges were developed during the creation of the PHIAI capability. In response to the lack of initial training data, I developed a generative AI text model using a Variational AutoEncoder, Seq2Seq architecture, and LSTM for the purpose of creating needed training data. This solution not only addressed the challenge of limited training data but also contributed to improvements in the algorithm's machine learning model complexity. To address the algorithm's low accuracies, I implemented a new stacking classifier that utilized the capabilities of several classifiers within one singular model. This resulted in a drastic improvement in algorithm accuracy, consistently achieving accuracies over 95% and significantly outperforming the algorithm's previously implemented model.

4. ANTICIPATED RESULTS

Although the PHIAI capability is still in development, the current results of PHIAI's algorithm demonstrate its effectiveness within real world applications. The capability is able to consistently identify potentially hazardous military subsystems with accuracies surpassing 95%. While this proves that the capability will be useful to System Safety Engineers when assessing technical systems, it also proves that the project will require additional training in order to maximize its potential in preserving user safety. Following my development of the PHIAI capability, the newly-designed algorithm utilizes more complex machine learning models and trains on an increased amount of data, improving model performance and algorithm accuracy by an estimated 20% from the previous implementation. This improvement significantly increased the reliability and effectiveness of the algorithm overall.

The results prove that the capability will be useful tool for System Safety Engineers. The PHIAI capability automates and streamlines the time-intenstive hazard identification process for System Safety Engineers. In doing so, System Safety Engineers will be able to focus more time on hazard analysis, hazard mitigation, and other important tasks. This will significantly improve the performance of System Safety Engineers by allowing them to be more productive in reducing the likelihood of injury/harm during complex technical military system use.

5. CONCLUSION

The PHIAI capability contributes significantly to governmental efforts focused on improving military system user safety by automating the potential hazard identification process for System Safety Engineers. The process is automated using an algorithm built around four central components involving several complex machine learning models and a graph-based database which stores military system design data. Using these components, the capability's algorithm is able to train and predict potential hazards without direct interaction from System Safety Engineers, streamlining the time-intensive task hazard identification process and providing System Safety Engineers more time to focus on other important matters. Although the PHIAI capability is still in development, its accurate and reliable results prove that the capability is useful and applicable to System Safety Engineers for preventing accidents and loss during military system use.

6. FUTURE WORK

Looking toward the future, the technology must be finished and deployed, tested to ensure effective operation, and improved to become applicable to systems in other fields and sectors. While the PHIAI capability achieves consistent accuracies surpassing 95%, the capability will need to be developed even further in order to make it more reliable to System Safety Engineering clients. Additionally, the capability will need to be tested rigorously to ensure that its algorithm works effectively and efficiently without negatively impacting the day-to-day operations of System Safety Engineers.

Finally, although the PHIAI capability is currently oriented toward System Safety Engineers and the military, proper research and improvements of the capability's usability, accuracy, and performance can make it applicable to any industry that promotes heavy interactions between humans and complex technical systems.

REFERENCES

Lawrence, R.G. (2022, August 17). Safety tech boosts productivity, contractors say. *ConstructionDive*. https://www.constructiondive.com/nws/safety-tech-boosts-productivitycontractors-say-/629826/

Nzongo, F. (2023, April 26). How can humans and intelligent computers work together? *Medium*. https://uxdesign.cc/how-can-humans-and-intelligent-computers-work-together-d349328ce270