SPACE ISAC

Artificial Intelligence & Machine Learning Community of Interest

MACHINE LEARNING SECURITY OPERATIONS

MLSecOps

by Max Spolaor, Ph.D. (The Aerospace Corporation), Trisha Miller (RS21), Michelle Archuleta (RS21) & Drew Wilson (Lockheed Martin)

Artificial Intelligence and Machine Learning (AI/ML) in Space

AI-enabled solutions are becoming increasingly important for securing operations and ensuring strategic advantages in the expanding and evolving space industry, both in the public and private sectors. As the use of AI in space applications grows, currently valued at around $2 billion, there is a need for more mature AI/ML solutions and standards (Defensescoop. 2023, March 13). Private space exploration has contributed to this growth, and governments are also investing more in AI-enabled technologies in a 21st century space race for AI-enabled solutions.

Agencies and organizations are prioritizing AI research and development efforts. However, AI-enabled space systems remain vulnerable to security threats across the AI/ML lifecycle – from initial data collection to deployment. The limitations and characteristics of the space environment are inconsistent with the processes and procedures currently used to secure AI, leading to vulnerabilities in AI-enabled technologies, primarily due to the lack of dedicated on-orbit systems and processes for securing and maintaining them.

The Space ISAC AI/ML Community of Interest (CoI) is focused on the intersection between AI technologies, cybersecurity, and space infrastructure. These technology areas are advancing rapidly, evolving the security landscape in parallel. The AI/ML CoI's mission is to provide thought leadership, guidance, and resources for improving the secure operations of space infrastructure through trustworthy AI technologies. The goal of this whitepaper is to establish Machine Learning Security Operations (MLSecOps) as a term of art and an implementation of Trustworthy AI that can be leveraged by the space community when developing and deploying secure AI-enabled systems.

The Origin of MLSecOps

The concept of MLSecOps was initially conceived as the logical reaction to the many technical challenges faced by enterprises to deploy AI-enabled solutions into production. Numerous surveys report that around 70% of AI/ML models never bridge the gap from the prototype stage to production. According to Sculley et al. (2015), the findings discussed in "Hidden Technical Debt in Machine Learning Systems" were initially investigated. The reason for this failure is rooted in the fact that a functioning AI-enabled solution requires the amalgamation of many codes, algorithm choices, hyperparameters, serving infrastructures, machine resource managements, data, and more. The reliance on data, which evolves as the underlying events change and as the data models themselves are refined, requires new techniques to facilitate development that are not sufficient for traditional deterministic code.

MLSecOps is analogous to the widely accepted Development Security Operations (DevSecOps) methodology, which merges Development Operations (DevOps) with Security Operations (SecOps) to deliver an Agile-style software development while integrating security into every aspect of the development process. Machine Learning Operations (MLOps) follow similar learning and methodologies to DevOps, with the crucial distinction that AI/ML processes observe a data-centric rather than code-centric philosophy. While DevSecOps is being adopted as best practice for space systems, there is a gap in addressing cybersecurity of AI-enabled technologies throughout their lifecycle when deployed in space systems. A clear place to incorporate security best practices into

AI-enabled systems is in the MLSecOps framework, however, MLOps and MLSecOps are early-stage concepts.

Like DevOps, MLOps is now a core function of ML engineering, focused on streamlining the process of taking ML models to production and then maintaining and monitoring them. MLOps is a collaborative function, often requiring data scientists, DevOps engineers, and IT personnel. MLOps builds on the concept of DevOps by employing continuous training along with Continuous Integration/Continuous Deployment (CI/CD) of models to mitigate the impacts of data drift and concept drift on the performance of models in production. MLOps consists of practices broadly associated with the following seven pipeline stages of ML model development and operations (see Figure 1):

1. Object Specification: define business problem, objectives, and key performance indicators.
2. Data Acquisition and Storage: identify and effectively store training and testing datasets.
3. Data Curation and Tracking: organize, integrate, and normalize data from various sources and multiple domains. Ensure data provenance. Standardize access and hooks for ML pipelines.
4. Model Training: build and train ML models. Experiment and tune models' performance.
5. Model Acceptance Testing: verify that ML models meet required/desired performance metrics. Ensure model quality, reproducibility, performance, explainability, etc.
6. Model Deployment: package, test, validate, and promote-to-production the ML models. Service continuous integration/ continuous deployment (CI/CD) infrastructure for production models.
7. Model Monitoring and Control: monitor performance once ML models are deployed. Ensure key performance indicators are satisfied. Retrain and push model updates according to results over time.

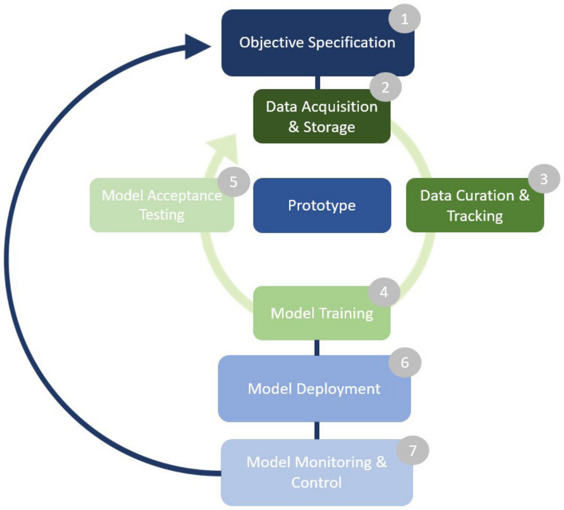

Figure 1: The Seven stages of the Machine Learning Operations (MLOps) life cycle.

To address the challenges and significant efforts needed to implement a full MLOps system, practitioners have defined a series of levels of implementation of MLOps to facilitate technology maturation. At Level 0, models are manually retrained and redeployed, whereas, at Level 4, the training and testing of models with a CI/CD pipeline is fully automated.

MLSecOps to Mitigate Novel Security Vulnerabilities

The adoption of MLOps methods and processes for AI-enabled space applications is still relatively early. Examples include an AI-enabled satellite maneuver detection technology from Booz Allen Hamilton, capable of deploying with a Modzy MLOps platform source and satellite image recognition technology from NASA's Implementation and Advanced Concepts Team (IMPACT), which is made flight-ready by deploying with MLOps components and workflows.

MLOps promises to accelerate the successful deployment of AI-enabled solutions in space, but AI/ML models and operations have unique security vulnerabilities. Vulnerabilities can manifest in AI/ML technologies throughout their lifecycle – from data collection and curation to inference. Correspondingly, adversaries are developing new tactics, techniques, and procedures (TTPs, see for example Aerospace SPARTA TTPs) to specifically exploit AI-enabled systems. MITRE's publication of the ATLAS Matrix – a knowledge base of adversary tactics, techniques, and case studies for machine learning systems, modeled after the MITRE ATT&CK framework, enables researchers to navigate the landscape of threats to machine learning systems and raise awareness of these threats in a familiar way to security researchers – highlights unique TTPs to exploit vulnerabilities in data, models, and broader AI systems. Furthermore, new vulnerabilities will be created with the evolution of autonomous space systems relying on AI-enabled technologies for operations or cybersecurity.

To ensure that AI-enabled technologies are both performant and resilient to cyber threats when deployed in space systems, we propose MLSecOps as a foundational concept for implementation of AI/ML in space. Aerospace has defined a Trusted AI Framework, which enables AI practitioners and stakeholders to evaluate the trustworthiness of an AI technology across their full lifecycle, from design and implementation through operation and maintenance of the technology (Slingerland et al., 2022). Implementation of MLOps is described as a "means through which trust is proven and maintained." MLSecOps takes this a step further by emphasizing the importance of security in the implementation of MLOps to enhance the proof of trust in deployed AI technologies and to ensure continued secure operations. Concepts for implementing secure AI deployments include both the engineering of AI solutions and a deployment approach which incorporates full membership in a security-focused ecosystem to actively identify and defend against advanced threats (see for example Aerospace SPARTA on-board intrusion detection and prevention countermeasures).

Recently, Space ISAC conducted the "Hera's Revenge" tabletop exercise to evaluate the effectiveness of different security postures under a Ground Station as a Service (GSaaS) cyber-attack scenario. This exercise highlighted to AI/ML CoI members how the use of AI technologies could enhance insights into space Cyber Threat Intelligence (CTI) data and accelerate decision-making in operational missions. This insight can potentially enhance the cyber-attack vector contextualization within the CTI data to increase the aggregation accuracy delivered to space operations. As AI technologies are adopted in space operations, like those exercised in Hera's Revenge, concurrent implementation of MLSecOps will provide best practices for the use of these technologies to establish and maintain trust throughout the ML lifecycle.

To remain competitive with adversaries, a comprehensive MLSecOps framework is essential for the United States to achieve space superiority. It provides real-time monitoring, analysis, and response to potential threats, ensuring the security and integrity of space systems. However, MLSecOps is a nascent field that still requires significant research investments to become as widely accepted and effective as DevSecOps. The Space ISAC AI/ML CoI aims to help bridge this gap by advancing the space community's awareness and knowledge of MLSecOps.

References

Defensescoop. (2023, March 13). Pentagon requesting more than $3B for AI, JADC2. https://defensescoop.com/2023/03/13/pentagon-requesting-more-than-3b-for-ai-jadc2/

Sculley, D & Holt, Gary & Golovin, Daniel & Davydov, Eugene & Phillips, Todd & Ebner, Dietmar & Chaudhary, Vinay & Young, Michael & Dennison, Dan. (2015). Hidden Technical Debt in Machine Learning Systems. NIPS. 2494-2502.

Slingerland, P., & Perry, D. (2021). Towards trusted AI in space systems. Aerospace Corporation. https://csps.aerospace.org/sites/default/files/2021-08/Slingerland_Perry_TrustedAI_20210719.pdf