Kyle DeProw

Active Top Secret | 🔗 LinkedIn | 📱 618-818-4583 | 🌐 kyle-deprow.github.io | ✉ kydepro@gmail.com | ⌥ GitHub Profile

Skills

- Python | C++ | C | Java | Rust | R | Lua | Matlab | Ada | PyTorch | TensorFlow | ONNX | CUDA | ROS | SQL | PostgreSQL | NoSQL | Git
- Pandas | Azure | AWS | CI/CD | Docker | Kubernetes | Kafka | Protobuf | Spark | Databricks | Edge Deployment | Time-Series Analysis
- Computer Vision | Multimodal Modeling | Anomaly Detection | VAE | RNN | CNN | BERT | YOLO | Diffusion | ViT | HuggingFace | LoRA

Experience

Senior Machine Learning Engineer — Boeing — *Berkeley, MO* **12/2021 - Current**

- **Program Management and Execution:** Conducted in-depth data analysis and experimentation, synthesizing complex findings from advanced ML techniques while collaborating with customers to align actionable strategies with business requirements, influencing a $16B portfolio in cybersecurity and AI safety domains using Azure, ElasticSearch, Spark, LLM finetuning, PyTorch, and Explainable AI.
- **ML Solution Development:** Led over $15M in R&D investments to advance AI techniques, building on SOTA algorithms GAN, AutoEncoders, Transformers (BERT, ViT) with a focus on novel solution designs, network architectures, and deployment/hardware optimizations for NLP, Computer Vision, and Time-Series Analysis problem domains.
- **Model Design and Development:** Drove innovation to pioneer $10M machine learning research and development effort to create a real-time behavioral monitor embedded in avionic platforms. Bayesian-based, custom VAE solution trained in distributed PyTorch executed in Azure Databricks and built on an ETL pipeline processing hundreds of GBs/hr using Spark, capable of detecting anomalies from expected norms in platform data and cross-correlating these events to rich log data stored in ElasticSearch within milliseconds.
- **Model Training and Evaluation:** Championed a rigorous statistical approach for training experimentation, fine-tuning and causal inference through confidence intervals, CasualML, and Bayesian optimization improving model F1 by 5-25%. Generated actionable insights for leadership by translating model evaluations into clear, data-driven narratives using visualization tools Matplotlib and Tableau.
- **Tailored Edge and Cloud Deployment:** Designed scalable, distributed data systems for cloud and edge use-cases. Leveraged Azure, PySpark, Databricks, and Kubernetes to achieve 99.9% uptime for critical workflows in cloud environments, while deploying optimized, low-latency models on edge devices using a PyTorch-to-ONNX-to-TensorRT pipeline for efficient resource utilization.
- **Model (Transformer) Fine-tuning:** Fine-tuned transformer-based LLM models using LoRA, achieving secure and scalable model adaptation for novel applications, such as translating unsafe C code into Python with 85% memory-safe performance.
- **Technical Leadership:** Managed and mentored a team of 12 multi-disciplined engineers while hosting enterprise-wide workshops to foster collaboration with cross-functional teams and combine outside domain knowledge, such as DevOps, IT, Cybersecurity.

Autonomy Engineer — *Hazelwood, MO* **02/2019 - 12/2021**

- **Autonomy / AI:** Designed and developed responsible and explainable AI agents for Aerospace Simulation (AFSIM) for optimal path-planning behaviors such as "detect and avoid" to support $1.3 billion effort using rule-based logic, A*, and finite-state machines in C++.
- **Optimal Path Planning with RL**: Designed and implemented reinforcement learning solutions for autonomous decision-making, leveraging architectures inspired by AlphaGo, including Monte Carlo Tree Search (MCTS) and deep neural networks. Enhanced traditional AI methods by integrating policy networks for strategic action selection and value networks for state evaluation, achieving a 40% improvement in planning efficiency of dynamic aerospace navigation and autonomous conflict resolution environments.
- **Autonomous Visual Perception:** Developed advanced perceptual solutions for autonomous aircraft, integrating fused sensor inputs (IFF, IMUs, Radar, Camera, LIDAR) with AI models (GNN, CNN, RL). Developed a high-precision visual tracking system, improving object detection accuracy by 15% over traditional techniques and achieving an overall 10% performance boost in dogfight simulations.
- **Model Experimentation and Optimization:** Performed data analysis utilizing core simulation experiments, including A/B testing, ablation studies, model cross-validation, and data refinement improving base performance (F1 and AUC) of models 10-50%.
- **DevSecOps and MLOps:** Implemented MLOps pipeline with IaC principles using Azure, Kubernetes, Docker and Ansible for cloud resources to reduce simulation setup time by 70% and standardize ETL deployment procedures among enterprise environments.
- **Continuous Integration/Deployment**: Gitlab pipeline Integration, pull requests, code reviews, stress testing, unit/integration testing.

Robotics and AI Researcher — Saint Louis University — *St. Louis, MO* **01/2017 - 02/2019**

- **Research and Publication:** Salaried position to lead and conduct projects, publish articles, and support conference presentations for research relevant to NSF Cyber-Human System programs in fields of Robotics, AI, and Machine Learning.
- **Supervised Learning:** Designed and developed systems facilitating the supervised learning of tactile and kinematic features on an anthropomorphic robotic arm using ROS, Tensorflow LSTM structures, and Pandas for Time-Series data-analysis.
- **Reinforcement Learning:** Explored Reinforcement Learning algorithms to guide robotic manipulation using ROS and Kalman Filters for state-space exploration in medical and rehabilitation contexts, achieving a 55% reduction in computational overhead in optimal scenarios.

Education

Masters of Science   **Saint Louis University**   *St. Louis, MO*   **01/2017 - 12/2022**

- Major in Engineering; Emphasis on Robotics and AI (GPA 3.85)

Bachelors of Science   **Southern Illinois University**   *Edwardsville, IL*   **08/2012 - 05/2016**

- Major in Mechanical Engineering, Minor in Mathematics (GPA 3.89)

Projects

Behavioral Analysis – DNN Enhanced Cyber Resiliency

- Led a $10M, three-year initiative to develop a large-scale, cloud based cyber anomaly detection service that ingested GBs/min/platform of operational telemetry data to create probabilistic operational profiles and flag real-time deviations.
- Cradle-to-grave project pioneered as an exploratory study against historic flight data in a Jupyter notebook which from this success was then scaled to a distributed cloud service with functioning ETL pipelines orchestrated in Databricks and PySpark.
- Deployed structured database leveraging Medallion architecture to collect, clean, and transform sensor data necessary for deep learning analysis. Cloud technology AKS used to provide Reliability, Scalability, and Maintainability for data consumers resulting in deployment reduction time by over 80%, scaled inference of service over 1000%, and team efficiency boost of 13% when addressing system faults.
- Model architecture includes Bayesian DNN structures built in PyTorch and achieving 75-90% F1 scores on abnormal custom test cases.
- AI Safety and explainability solutions, such as saliency maps and feature attribution, inspired by relevant literature and frameworks like Anthropic's Responsible Scaling Policy additionally provided in solution to inform CONOPS for autonomous systems, enabling customers to trust AI-driven decision-making processes and ensure safe scaling of models to meet customer operational needs.

Binary Analysis – DNN Similarity Comparison Approach

- Applied foundation models and custom architectures, including transformer and multimodal approaches, to fine-tune and train models respectively for analyzing disassembled binary code. Developed benchmarks and demonstrated a 90% F1 performance. This work resulted in wide adoption within cyber red teams who, on average, identify 3 code risks per test event saving the company $Ms in mitigated escape defects and exploitable code.
- Functional blocks included: RoBERTa LLM for interpreting disassembled microcode, ResNet classifier for ingesting file binary entropy to identify architecture/compiler environment, and an Attention later to cross-correlate patterns with known binary code.
- Constructed dataset using 10 open-source utilities (linux-core) built on 2 compilers, 6 optimization levels, 3 architectures, and 2 word sizes to obtain approximately 600k binary functions and 14M basic blocks.

Cyber Patriot – Boeing/AFA Youth Outreach

- Implemented Boeing's Data Science Module for 2024 Cyber Patriot activity which focused on behavioral-based anomaly detection of simulated flying platforms.
- Docker-based solution, deployed in Ansible allowed student participants to actively query a platform's avionic and guidance data through REST API (FastAPI) calls.
- Challenge exercises involved teams being asked to build basic data statistics and distributions to flag events of "malicious behavior" (such as spoofing the GPS positional data of a platform) injected into the simulation to purposely exceed expected data norms.

GrandM – Dynamic Chess Engine

- Project designed and built to prototype dynamic chess-playing AI utilizing advanced Finite State Machines to power personal game development projects.
- Objective is to develop a non-traditional AI (non-Stockfish or equivalent) capable of utilizing potentially modified chess pieces and chess boards to create an engaging and performant opponent for the player.
- MiniMax, Negamax, Alpha-beta pruning, Monte-Carlo Tree Search algorithms deployed to explore the trade-offs between optimal gameplay, computation cost, future positioning, and overall player satisfaction with the AI.

Certifications

- Software Engineering and Architecture: **Saint Louis University**
- Data Pipelines and Orchestration with Apache Airflow: **Saint Louis University**
- AWS Academy Solutions Architect Professional Certificate: **Saint Louis University (Expected 2025)**