Hoang (Joe) Do

703-397-6188 | joedo2910@vt.edu | linkedin.com/in/hoanglehuydo/ | github.com/notjoedo | joedo.dev

EDUCATION

Virginia Tech   Blacksburg, VA
B.S. in Computational Modeling and Data Analytics, Minor in Computer Science, Math   *Aug. 2023 – May 2027*
Concentration: Cybersecurity & Cryptography   **GPA**: 3.66/4

TECHNICAL SKILLS

Software Skills: Python, Java, JavaScript/TypeScript, C, C#, HTML/CSS, R, SQL, React, Next.js, FastAPI, Django, Flask, Robotic Process Automation (openRPA, UIPath), Parallel Computing (CUDA, OpenMP), Splunk (SPL)
Developer Tools: Git, Docker, AWS, Azure, PostgreSQL, Vercel, Figma

EXPERIENCE

AI Integration & Full-Stack Software Engineering Intern   May 2025 – July 2025
Centennial Technologies   *Leesburg, VA*
- Architected and deployed a **Retrieval-Augmented Generation (RAG)** system for Agentic-AI platform, utilizing Azure services **(Blob Storage, AI Foundry, Functions, VM)** to automate data sourcing, enabling users to query FDA approval document data in under 3 seconds.
- Engineered a hybrid Agentic-AI solution by developing a **Robotic Process Automation (RPA)** bot using **Optical Character Recognition (OCR)** and **Azure services** to automate data entry/extraction for XFA files, eliminating manual bottlenecks and enabling Agentic-AIs to autonomously execute end-to-end workflows by 90%.
- Designed UI/UX from wireframes and developed a modern frontend **(React, TypeScript, Tailwind)** that provided an intuitive and accessible interface, passing WCAG tests for Agentic-AI document processing system.

Lead SWE Undergraduate Researcher   Aug 2024 – Present
Virginia Tech GrayUR's Phased Array Microphonics Team   *Blacksburg, VA*
- Directed a 5-member cross-disciplinary team, serving as the technical leader for the integration of custom algorithms with phased array hardware and utilizing **Trello** to manage dependencies and track complex project deliverables.
- Enhanced project's analytical capabilities by researching, developing, and integrating a **custom PyTorch model**, improving data prediction accuracy by 55% and automated a previously manual analysis process.
- Established model serving and data management protocols **(SQL/NoSQL)** to ensure the scalability and reliability of the PyTorch prediction service, drastically reducing latency and facilitating real-time data access.

Splunk Dashboard Data Analytics Developer Intern   Aug 2024 – Present
Virginia Tech Division of Information Technology   *Blacksburg, VA*
- Enhance system monitoring by translating stakeholder needs into **15+ analytical dashboards using Splunk SPL**, which reduced the identification time for critical security events by 90%.
- Detect system vulnerabilities in real-time by developing a smart alert system in Splunk that successfully **identified over 100 critical errors and security breaches daily**, leading to 50% decreased incident response time.
- Optimized data ingestion performance by architecting new log forwarding configurations and data validations across 10+ data sources, **resulting in a 30% reduction in search latency and improved system reliability.**

PROJECTS

Atlas - Marriott 2025 CodeFest Winner | *Vite, Gemini API, Flask, MongoDB*   October 2025
- Architected full-stack solutions using **Flask REST APIs, Vite, and MongoDB Atlas** for documents + vector search for cached analytics.
- Developed an AI supported ideation and creation pipeline using **Google Gemini LLMs, UploadThing image handling, and prompt-based itinerary generation**.

GiggleBytes - 1st Place Data Competition Winner | *Python, Scikit-learn, TensorFlow*   November 2025
- Architected a full-stack solution that processed and cleaned complex datasets **(NumPy, Pandas, tidyverse)** to generate insightful scores, delivering via a custom interactive map visualization **(Typescript, Vite)** for actionable comparison.
- Developed and validated a ML model **(Scikit-learn/TensorFlow)** to identify key predictive variables, integrating into data processing pipeline for competition's actionable recommendations.