

# Jacob Furtaw

[jfurtaw97@gmail.com](mailto:jfurtaw97@gmail.com) | 410-533-7663 | [My LinkedIn](#) | [My Website](#)

---

## Professional Summary

Innovative Machine Learning Research Engineer with an expertise in Natural Language Processing and Computer Vision. Highly skilled in developing cutting-edge AI solutions with large language models (LLMs) and multi-modal systems. I am experienced in leading technical teams and driving projects from ideation to deployment. I am passionate about solving complex problems and pushing AI boundaries. Please explore my website to view my full [virtual resume](#) and [recent projects](#).

## Skills

**AI/ML & Data Science:** PyTorch, TensorFlow, Transformers, HuggingFace, Langchain, Llama-Index, Scikit-Learn, Pandas, NumPy, Matplotlib, Keras, Ollama, Cuda, Accelerate, **Retrieval-Augmented Generation (RAG)**, and Agentic RAG

**Programming Languages:** Python, JavaScript (Node.js), and CSS; **Familiar With:** C++ and Java

**Tools & Platforms:** Git, Docker, Kubernetes, Jupyter, Conda, PyCharm, VS Code, Linux, Windows, Nvidia NIMS

## Professional Experience

**Machine Learning Research Engineer** | SurgePoint Software | Hybrid

- Developing an innovative AI Assistant using open-source LLMs like Qwen 2.5 7B, Mistral-Nemo 12B, and Llama 3.1 8B
- Utilizing **data engineering** skills to reduce 200 million lines of unstructured data into a 13-million-line structured dataset
- Designing, engineering, and testing an advanced **Retrieval-Augmented Generation** pipeline that feeds any LLM the newest data from the university website to increase semantic relevance scores and combat model hallucinations
- Leveraging **prompt engineering** skills to ensure the prompt was well-structured and the model could answer any query

**Advanced Repair Agent** | Geek Squad | On-Site

- Recognized with a Letter of Recommendation for designing and implementing operational improvements, resulting in a 150% increase in productivity and customer satisfaction.
- Consistently ranked in the top 3% of all Advanced Repair Agents in the District and Region
- Completed tens of thousands of hardware and software repairs on various devices and operating systems

## Projects

### Chat RAG

- Developed a versatile RAG (Retrieval-Augmented Generation) chatbot with a Gradio user interface, supporting multiple local models from Ollama and HuggingFace and cloud-based from OpenAI, Anthropic, and NVIDIA NIMS
- Engineered a modular architecture with utilities for model management, chat functionality, and UI interactions, featuring dynamic model switching, custom prompt integration, model parameter tuning, and quantization options
- Designed flexible data ingestion from diverse sources (local files, GitHub repositories, and databases)

### **Automatic Identification of Equivalent Mutants using an ASTNN**

- Excelled as a member of a diverse Scrum Team, engaged in sprint planning, daily standups, and sprint reviews
- Investigated and implemented the use of a transformer-based model for the task of binary classification
- Optimized data preprocessing, added and tuned hyperparameters, and increased the model's overall performance
- Condensed research and outcome into a comprehensive formal report my team presented at a campus research fair

### **Image Alter**

- Developed a diffusers pipeline for text-to-image and image-to-image generation with a user-friendly Gradio UI
- Implemented customizable model parameters to give users more control over the images they generate
- Optimized model loading and inference using PyTorch and CUDA for efficient GPU utilization

## Education

### **Bachelor of Science in Computer Science, Software Engineering Concentration**

Towson University, Towson, MD

- **Coursework:** Artificial Intelligence, Database Management Systems, Software Quality Testing & Assurance, Data Structures and Algorithms, Requirements Analysis and Modeling, Software Engineering, Software Project Practicum

### **Associate of Art in Computer Science**

Howard Community College, Columbia, MD