

Jacob Furtaw

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

Intro

I am an innovative Machine Learning Research Engineer specializing in Natural Language Processing (NLP), seeking opportunities to advance AI technology. I excel in both independent and collaborative settings, driving projects from ideation to deployment. I am highly experienced in developing cutting-edge AI Assistants and implementing advanced NLP techniques like RAG for versatile, high-impact solutions. My expertise spans data preprocessing, synthetic data creation, model training, custom AI assistant development, RAG pipeline creation, and deploying scalable Machine Learning solutions. 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

Machine Learning/Data Science: PyTorch, TensorFlow, Scikit-Learn, **Transformers**, Hugging Face, **Langchain**, **Llama-Index**, Gradio, Pandas, NumPy, Data Preprocessing, Synthetic Data Creation, **Retrieval-Augmented Generation**

Programming Languages: Python, JavaScript (Node.js), CSS, C++, Java

Tools & Platforms: Git, **Docker**, Linux, Windows, Jupyter, PyCharm, VS Code, Nvidia AI Workbench

Work Experience

Machine Learning Research Engineer | SurgePoint Software

Aug 2023-Present

- Developed an innovative AI Assistant using open-source LLMs like Mistral 7B, Mistral-Nemo 12B, and Llama 3.1 8B
- Utilized self-taught **data engineering** skills to craft a 13-million-line dataset of web-scraped data into a structured JSON file
- Implemented **Retrieval-Augmented Generation** to increase semantic relevance scores and combat model hallucinations
- Leveraged **prompt engineering** skills to ensure the prompt was well-structured and the model could answer any user's query
- Worked with a small, diverse, and time constrained group of individuals at a startup in a remote setting

Advanced Repair Agent | Geek Squad

Mar 2022-Present

- Earned a Letter of Recommendation from Upper Management for designing and helping to implement a plan to revamp our locations day to day operations, increasing overall productivity by 150% and customer satisfaction
- Executing tens of thousands of Software and Hardware fixes on desktops, laptops, tablets, and cell phones
- Receiving an award for Top Performing ARA at our location every quarter and was in the top 3% in our District and Region

Projects

Chat RAG

- Developing an advanced, chat-based coding assistant using Retrieval-Augmented Generation technology, integrating multiple language model providers (HuggingFace, Ollama, OpenAI), and implementing a user-friendly Gradio UI.
- Engineering a flexible system with dynamic model switching, custom embedding integration, and efficient memory management, showcasing proficiency in PyTorch, Llama Index, and state-of-the-art language models.
- Implementing key features including context-aware document retrieval, streamed responses for real-time interaction, and quantization options, demonstrating strong skills in ML engineering and AI application design.

Automatic Identification of Equivalent Mutants

Towson University, Towson, MD

- Excelled as a member of a 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 overall performance
- Condensed research and outcome into a comprehensive formal report my team presented at a campus research fair

Education

Bachelor of Science in Computer Science, Software Engineering Concentration

Jan 2022-Dec 2023

Towson University, Towson, MD

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

Associate of Science in Computer Science

Aug 2019-Dec 2021

