

Aishwarya Hastak

Bloomfield Hills, MI

(930) 215-5327 | ahastak@iu.edu | [Portfolio](#) | [LinkedIn](#) | [Github](#) | [Medium](#)

Summary

Data Scientist with 2+ years of experience in building optimized Data and Machine Learning pipelines for cloud and hardware. Expertise in predictive analytics, generative AI, LLMs, Deep Learning, Model optimization, and cloud architecture to develop scalable and high-performance solutions.

Skills and Certifications

ML/LLM Frameworks	TensorFlow, PyTorch, NumPy, Pandas, Matplotlib, Scikit-learn, Transformers, OpenAI, NLTK, PySpark
Languages and Databases	Python, SQL, JavaScript, MySQL, PostgreSQL, MongoDB
Tools	AWS, Spark, Hadoop, PowerBI, HuggingFace, MLFlow, Git, Github, Docker
Certifications	AWS Certified Solutions Architect Associate (SAA-03)

Work Experience

AI/ML Collaborator Retrieval Augmented Generation Langchain AWS	Bloomfield, MI
Concept Reply	October 2024 - Present
<ul style="list-style-type: none">Developed a RAG chatbot with Meta's Llama 3.2 8B model in AWS, using LangChain and CoT prompt engineering to enhance response accuracy.Enhanced search retrieval accuracy by 20% with a FAISS vector store using metadata tags and dynamic similarity search with scoringReduced response time by ~50 seconds for a resource-constrained application through model quantization, sharding, and multi-threading.	
Data Science Research Associate Prompt Engineering Javascript LLM	Bloomington, IN
Observatory on Social Media, Indiana University	December 2023 - Present
<ul style="list-style-type: none">Developed an AI agent using GPT 4o for the data visualization tool Helios-Web, enabling interactivity for network manipulation and exploration for large networks with over 100,000 nodes. Improved model response accuracy by 60% through extensive prompt testing.	
Machine Learning Research Intern PyTorch GAN Qiskit Linux	Bloomington, IN
Indiana University Bloomington	Jan 2024 - Aug 2024
<ul style="list-style-type: none">Co-authored a research paper on developing a novel <i>LSTM Quantum GAN</i> architecture to generate high-resolution images with 5 times reduction in qubit counts and 12 times reduction in two-qubit gates, achieving generator convergence within 200 epochs with only 5 qubits.Analyzed parallel compute performance of transformer architectures for upto 128 TPUs, using DNN accelerator simulation tools like SCALE-SIM.	
Associate Software Engineer Python SQL AWS Agile ETL	Pune, India
Accenture	June 2021 - July 2022
<ul style="list-style-type: none">Implemented 7+ data pipelines to ingest and transform real-time usage and monitoring data from smart electric meters, utilizing AWS Kinesis and Lambda functions to handle data from around 400,000 residential and commercial customers.Enhanced data querying efficiency by indexing and partitioning data based on time (daily and monthly). Used AWS Glue jobs to transform CSV files to Parquet format, resulting in a 2500 ms improvement in query performance through advanced data compression techniques.	
Data Engineer Intern MongoDB Python Hadoop MapReduce	Pune, India
Amar Ujala Web Services Pvt. Ltd.	June 2020 - July 2020
<ul style="list-style-type: none">Designed an ETL pipeline to ingest from MongoDB and aggregate data using MapReduce for 4 subsidiary e-newspaper websites of AmarUjala.Analyzed user behaviors and content preferences in Python, resulting in an 8% increase in active users through targeted content curation.	

Projects

Implementing LoRA from Scratch for Fine-Tuning Transformers LoRA LLM PyTorch	
<ul style="list-style-type: none">Fine-tuned DistilBERT model on the Emotions dataset from Hugging Face using Low-Rank Adaptation (LoRA), achieving a 93% accuracy.Utilized rank-decomposition with a rank of 1, reducing the model parameter size by 99%, significantly speeding up the training process.	
Ted Talks Retrieval-Augmented Application Python PostgreSQL Transformers ElasticSearch Docker Grafana	
<ul style="list-style-type: none">Engineered a Retrieval-Augmented Generation (RAG) system for efficient Ted Talks Q&A, integrating the Mistral 7B Instruct model and Elastic Search for advanced response generation and retrieval achieving a 0.87 Hit Rate and 0.85 Mean Reciprocal Rank (MRR).Utilized vector embedding using SentenceTransformers and document re-ranking to improve model performance by 12%Monitored the pipeline using user feedback tracked in a PostgreSQL database and visualized on a Grafana dashboard.	
NeuroDetect: MRI-Based Alzheimer's Disease Detection PyTorch DNN Sagemaker MLFlow	
<ul style="list-style-type: none">Built an end-to-end data-centric image classification DNN model to predict Alzheimer's disease, achieving a F1 score of 83%..Applied data augmentation techniques to balance the skewness in the dataset and gain a 11% increase in accuracy.	

Education

Indiana University Bloomington	Bloomington, IN, US
Master of Science in Data Science, CGPA - 3.63/4.00	May 2024
NMIMS University	Mumbai, India
Bachelor of Science in Computer Science, Meritorious Student with Distinction, CGPA-3.51/4.00	May 2021

Publications

- Co-authored a paper titled "*LSTM-QGAN: Scalable NISQ Generative Adversarial Network*" published in ICASSP, Sep 2024