Ashik Syed Shaffiullah

https://ashikshafi08.github.io/

Email : ashiksyedshaffiullah@gmail.com
Mobile : +1-312-900-7691

Experience

Founding Machine Learning Engineer February 2024 - Present
Nakamoto LLC *Chicago, IL*

- Engineered an AI-powered autonomous bug-fixing system with LlamaIndex-based RAG and LLM-driven patch generation, achieving a 32% success rate on the Princeton SWE Benchmark, surpassing baseline models by 15%.
- Leveraged pre-trained Qwen and DeepSeek models with FAISS-based retrieval augmentation for function writing systems, increasing code inference accuracy by 20% on BigCodeBench while reducing inference costs.
- Designed a fine-tuning framework for mathematical reasoning LLMs with custom loss functions and specialized MoE routing, boosting logical inference and answer accuracy across complex math tasks.
- Achieved top AI miner ranking on Bittensor through RAG-enhanced subnet deployment, earning ˜$200K/month by optimizing model performance with sparse attention mechanisms.
- Reduced LLM inference latency by 60% through model quantization, distributed model parallelism, and NGINX-based load balancing for efficient inference scaling
- Designed a full-stack mining solution with automated scripts and monitoring dashboard using Python, Docker, and Kubernetes for decentralized AI networks, reducing operational costs by 80% and generating $20K in three days.

Machine Learning Intern March 2022 - August 2022
NeuralMetrics.ai *Denver, CO*

- Built a Classification Engine using Python and AWS Lambda, enhancing business categorization accuracy by 25% through optimized NAICS/SIC code identification.
- Refined XGBoost models, increasing prediction accuracy by 20% and reducing processing time by 30% for improved risk assessment efficiency.
- Developed AI-driven workflows with real-time adaptive learning, elevating model accuracy by 35% within three months through continuous feedback integration.

Technical Skills and Certifications

Languages: Python (Pandas, NumPy, TensorFlow, PyTorch), SQL, Java, Jax
Frameworks: Sklearn, TensorFlow, PyTorch, Spark, Flask, FastAPI, Django, HuggingFace, LangChain
Tools: Docker, Kubernetes, Azure ML, Git, Jenkins, MySQL, MongoDB
Practices: RAG, LLMs, Prompt Engineering, CI/CD, RESTful APIs, A/B Testing, MLFlow
Certifications: Google Certified TensorFlow Developer

Education

Illinois Institute Of Technology August 2022 - May 2024
Master of Computer Science *Chicago, IL*

Panimalar Engineering College August 2018 - June 2022
Bachelors of Engineering in Electronics and Communications *Chennai, India*

Projects

Weave Framework – Synthetic Data Generator for Finetuning LLMs | *Python, PyTorch, Hugging Face*

- Architected a production-ready Python framework for AI-powered synthetic data generation, generating **1M+ high-quality synthetic samples** for LLM fine-tuning across various domains.
- Engineered specialized "noisers" for context-aware data augmentation, improving language model robustness and increasing synthetic dataset diversity by **30%**.
- Implemented intelligent dataset management tools for merging synthetic and real data, reducing preprocessing time by **40%** and ensuring high-quality data integrity with automated validation.

Open Source Contribution | *Python, PyTorch, PyTorch Lightning*

- Patched a critical double conversion bug in Bittensor's stake swap, preventing stake amount inflation (e.g., an incorrect 21.5 billion conversion from 21.5369).
- Refactored PyTorch Vision's LAMB optimizer, accelerating training speed by 15% on large datasets.
- Architected the Mosaic Commune subnet's image generation pipeline, improving speed and reliability.
- Revamped the leaderboard system for Synthia Subnet on Commune, slashing load time from over 50 seconds to under 5 seconds for 500 daily users using aiohttp and asyncio.

PubMed 200k RCT Implementation | *NLP, TensorFlow, BiLSTM*

- Fine tuned a deep learning model to classify sentences in medical abstracts into categories such as BACKGROUND, METHODS, and RESULTS, achieving **87.31%** accuracy.
- Implemented a hybrid approach combining token embeddings, character embeddings, and positional information to enhance classification performance.
- Optimized data preprocessing and conducted hyperparameter tuning to improve model efficiency and generalization.