Shriyansh Singh

+1-930-333-5141 | shriyansh.singh24@gmail.com | LinkedIn

SUMMARY

Machine Learning Scientist with expertise in **deep learning**, **LLMs**, and **natural language processing**. Experienced in developing **production-quality**, **scalable ML solutions** using **PyTorch** and **TensorFlow**.

PROFESSIONAL EXPERIENCE

Machine Learning Engineer *April 2024 - Dec 2024*
Hyphenova AI *Los Angeles, CA*

- **Designed** and **implemented** multiple **deep learning models** using **PyTorch** and **TensorFlow** that improved fraud detection accuracy by 32% while reducing false positives by 18%
- **Developed** and **fine-tuned** a domain-specific **LLM** using **LangChain** and **RAG** techniques, **collaborating closely** with compliance teams to ensure solutions met regulatory requirements
- **Architected** a **cloud-native** solution using **vector databases** deployed via **Kubernetes** to enable semantic search across 3M+ financial documents
- **Established** comprehensive **CI/CD pipelines** with **robust unit tests** to ensure reliable deployment of ML models, reducing production incidents by 63%

ML Research Intern *May 2022 - Oct 2022*
Enterprise Financial Technologies *Mumbai, India*

- **Implemented reinforcement learning models** for automated trading strategies, **collaborating closely** with trading desk teams to align models with business objectives
- **Designed experimental frameworks** with **intrinsic and extrinsic metrics** aligned with business KPIs to evaluate model performance in production scenarios
- **Optimized ML pipelines** for distributed training on **AWS** and **GCP** infrastructure, reducing model training time by 68% while maintaining accuracy
- **Developed production-quality code** with comprehensive error handling, logging, and monitoring to ensure reliable operation in business-critical environments

KEY PROJECTS

Financial LLM Assistant | *PyTorch, LangChain, LangGraph, Vector DB, RAG, AWS SageMaker* *Sep 2024 – Dec 2024*

- **Developed** a specialized **LLM application** with **production-quality code** including extensive error handling, monitoring, and **CI/CD integration**
- **Fine-tuned** open-source **transformer models** with **continuous A/B testing** to optimize performance based on business metrics
- **Collaborated** with compliance and legal teams to ensure the system adhered to regulatory requirements while maintaining high utility for end users

Distributed ML Training Framework | *PyTorch, AWS, GCP, Docker, Kubernetes* *May 2024 – Aug 2024*

- **Architected** a scalable framework for **distributed training** deployed on **cloud infrastructure** with automated scaling capabilities
- **Implemented** data parallelism and model sharding techniques with **extensive debugging tools** to ensure reproducible training results

TECHNICAL SKILLS

Machine Learning: Deep Learning, **Reinforcement Learning**, **LLMs**, **Natural Language Processing**, **Computer Vision**, Time Series Analysis
ML Frameworks: PyTorch, **TensorFlow**, **Scikit-learn**, **Keras**, **Hugging Face Transformers**
LLM Ecosystem: LangChain, **LangGraph**, **Vector Databases**, **RAG**, **Prompt Engineering**, **Fine-tuning**
Programming: Python, **SQL**, **NumPy**, **Pandas**, **SciPy**, R, CUDA, C++
MLOps & DevOps: CI/CD Pipelines, **Docker**, **Kubernetes**, **Unit Testing**, **Git**, **Production Debugging**, Jenkins
Cloud & Big Data: AWS SageMaker, **GCP Vertex AI**, **Distributed Computing**, Apache Spark, Hadoop, Data Pipelines

EDUCATION

Indiana University Bloomington *Aug 2023 – May 2025*
Master of Science in Data Science *Indiana, United States*

- Specialization: Machine Learning and Artificial Intelligence
- Relevant Coursework: Deep Learning Architectures, Natural Language Processing, Reinforcement Learning, Cloud Computing for ML, Software Engineering for ML, Statistical Machine Learning, Advanced Mathematics for ML
- Research Focus: Applications of Large Language Models in Financial Services