PRABHAKAR ELAVALA

Manchester, NH,03103 · +1 8573980726 · prabhakarpr554@gmail.com · LinkedIn profile · GitHub Profile

Professional Summary

AI/ML Engineer with hands-on experience in LLMs, MLOps, and full-stack AI applications. Certified in Azure AI & Databricks ML, with proven results reducing latency by 60% and deploying pipelines on datasets of 600K+ records

Work Experience

Data Science Intern   **April 2025 – July 2025**

CommandL | Remote

- Analyzed **10K+ learner sessions** to detect unclear instructions, improving content clarity & reducing confusion
- Audited **AI-generated instructional data** for tone, grammar, and logical flow, raising **QA pass rates by 25%**
- Produced **LLM performance reports** that guided prompt updates, reducing **content revision cycles by 20%**

Head of Chatbot Development (Volunteer)   *January 2025 – April 2025*

CareEscapes AI | Boston,USA

- Designed and deployed a **healthcare chatbot (Azure OpenAI + LangGraph)**, enabling **smart & easy navigation**
- Built scalable **WebSocket + Redis infrastructure**, cutting latency by **60%** while supporting **1,000+ users**
- Integrated **FastAPI, PostgreSQL, and React** into multi-agent pipelines, improving uptime & seamless UX
- Eliminated **90% of timeout errors** with progressive backoff, health checks, and automated session cleanup

Projects

Git-Guide - An AI Tutor for GitHub Repositories

- Built a **full-stack AI platform (Next.js, FastAPI, PostgreSQL)** that converts repos into AI-guided learning journeys
- Delivered a **context-aware AI assistant** that adapts tasks to user skill level, boosting learning efficiency
- Optimized **repo parsing & push verification** with GitHub API, reducing processing latency by **40%**

Song Popularity Prediction – End-to-End ML System

- Engineered **13-step ZenML pipeline with MLflow** to process **600K+ tracks** via smart sampling (5× faster)
- Extracted **18 audio features with Essentia** and tuned 5 ML models (XGBoost, LightGBM), achieving **R^2 = 0.51**
- Deployed via **Gradio + FastAPI backend** on Hugging Face, enabling **real-time audio-to-prediction inference**

Custom GPT Implementation from Scratch

- Built a **transformer-based language model from scratch**, showcasing mastery of deep learning fundamentals
- Implemented **multi-head attention, positional encoding, custom tokenization** for character-level modeling
- Accelerated training with **CUDA GPU optimization**, boosting throughput and enabling large-corpus training

Technical Skills

Machine Learning & AI: GPT-4, Azure OpenAI, Hugging Face, LangChain, MLflow, ZenML, scikit-learn, PyTorch
Big Data & Distributed Systems: Spark, Hadoop, Kafka, Dask, Flink, Airflow, Databricks, BigQuery, Snowflake
Data Engineering & Pipelines: SQL, ETL, Data Cleaning, Feature Engineering, Statistical Analysis, A/B Testing
Deployment & Infrastructure: FastAPI, REST APIs, Docker, Git/GitHub, Redis, Hugging Face Spaces, Azure Cloud

Education

M.S - Informatics (Concentration: Analytics)   **January 2024 – December 2025**
Northeastern University, Boston, MA   **GPA: 3.8**
B.Tech - Electronics and Communication Engineering   **August 2019 – June 2023**
Sree Nidhi Institute of Science and Technology (SNIST), Hyderabad, India

Certifications

Databricks Certified ML Engineer Associate: Proficiently applied advanced MLOps techniques on Databricks
Microsoft Certified: Azure AI Engineer Associate: Skillful in designing, evaluating, and scaling AI solutions