Gnana Prasuna Nimeesha Vakacharla

San Jose, CA | nimeeshav02@gmail.com | (203) 809-7684 | Portfolio | linkedin.com/in/nimeesha-vakacharla-811340289 | github.com/Nimeesha-Vakacharla

Education

San Jose State University — San Jose, CA
Master's in Data Analytics — Jan 2024–Dec 2025
Coursework: Probability Distributions, Database Management Systems, Big Data Technologies, Distributed Systems, Agentic AI

Technical Skills

Programming Languages: Python, JavaScript, R, C, SQL, NoSQL, HTML, CSS, Perl, NodeJS
Machine Learning & Statistical Methods: Scikit-learn, TensorFlow, PyTorch, XGBoost, LightGBM, OpenCV, t-test, Chi-square
Cloud & Data Engineering: AWS (S3, EC2, Lambda), GCP, Azure, Py Spark, Hadoop, Hive, Docker, Kubernetes, CI/CD, Databricks
Data Analysis & Visualization: Pandas, NumPy, Matplotlib, Seaborn, Tableau, PowerBI, AWS QuickSight, Plotly, MS Excel
Deep Learning & Generative AI: CNN, BiLSTM, GANs, Diffusion Models, Transformers, PEFT, LoRA, NLP, LLM, RAG

Experience

Gen AI Engineer — May 2025 – Present
Re-Vive, Texas

- Streamlined end-to-end **Gen AI-powered dashboard solution** for enterprise process analytics, **implementing comprehensive OpenAI integration** to deliver intelligent storytelling capabilities reducing $8M operational costs
- **Deciphered AI-driven insight summarization engine** processing 13K+ service requests, optimizing over user experience to transform complex banking workflows into actionable narratives for 150+ analysts
- Accelerated production-ready UI/dashboard with dual-mode Gen AI capabilities, **driving measurable impact with 20+ intelligent instruction types** and real-time AI-fueled process observability across US/Canada regions

Data Scientist — May 2024 – July 2024
iQuadra, Nellore

- **Optimized end-to-end enterprise book recommendation system** serving 3M+ users, **exploring 5.6GB dataset** using Hadoop/PySpark, **delivering 78% customer engagement improvement** through bias for action
- **Invented 6 MapReduce automation solutions** analyzing user behavior patterns, obsessing over system performance to **reduce processing time by 65%** while maintaining uncompromising standards for enterprise deployment
- Amplified production-ready Flask application with Plotly business intelligence visualizations, **delivering measurable business impact through 42% accuracy improvements and 60% faster response times** using frugal algorithms

Academic Projects

Interview Prep-AI | Python, RAG, Vector DB, Airflow, GCP, Mistral-7B, Llama-4

- **Built AI-enabled interview simulation platform** using fine-tuned LLMs (Mistral-7B, Llama-4), diving deep into user needs to **achieve 90% question accuracy** and **85% evaluation precision**
- Architected scalable data pipeline processing 4,500+ interview datasets with Airflow/GCP, **built to handle 10x future capacity** while **delivering 40% faster preprocessing** through **resource-efficient automation solutions**
- Deployed production-ready platform **supporting 100+ concurrent users** with vector database optimization, taking full ownership of performance metrics with **97% response uniqueness** and **lightning-fast less than 5 second query performance**

DIS-EASIFY | Python, Django, Azure, Scikit-Learn, TensorFlow, Docker, Kubernetes, MLOps

- Pioneered AI-supported healthcare platform using TensorFlow/Django, **prioritizing global patient outcomes with 97.6%-100%** accuracy across heart disease, diabetes, pneumonia, breast cancer detection **for underserved populations**
- Engineered scalable Azure infrastructure with Docker/Kubernetes CI/CD pipelines, prioritizing seamless user experience for global scale serving 1000+ concurrent users with **frugal less than 2s response times** and HIPAA-compliant security
- **Delivered 60% healthcare accessibility improvement** through probability-based assessments, rigorous analysis of 5+ datasets with ML algorithms, **insisting on highest standards to reduce false positives 85% and deliver reliable patient outcomes**

Cloud Analytics on New York Times Data | Python, Apache Airflow, GCP, BigQuery, Power BI

- Automated ELT pipeline extracting 500K+ JSON records from NYT APIs using Python/Airflow, **optimizing data reliability mechanisms to achieve 99.9% uptime** while thinking big for 1TB+ historical processing
- Created GCP data warehouse with BigQuery star schema and DBT transformations, maintaining highest performance standards to **deliver 40% query improvement** while maintaining zero data corruption across 1,000+ daily records
- Synthesized ML forecasting platform with polynomial regression **achieving 85% prediction accuracy**, earning stakeholder trust through Power BI business intelligence dashboards showing 11.7% **customer-impacting growth** across KPIs

Leadership and Achievements

- **Delivered top 8%** in NVIDIA Hackathon, **outperforming 200+ teams** through strategic **innovative technical innovation**
- pouvselectfont **Coordinated 500+ student operations** as Department Representative, **streamlining cross-departmental processes**
- **Mastered 6+ years** of Veena and Carnatic music performance, **demonstrating disciplined learning and precision** in cultural events