Vamsi Agnihotram

San Jose, CA | (669)-400-9964 | LinkedIn | Github Profile | agnihotramvamsi@gmail.com

Technical Skills

Languages: Python, C++, Java, SQL, Bash, JavaScript
AI & Machine Learning: TensorFlow, PyTorch, Keras, ONNX, BERT, MLflow, Forecasting Models, CNN-ANN
NLP: SpaCy, Hugging Face Transformers, Prompt Engineering, Text Classification, Sentiment Analysis
Systems & Integration: ROS, gRPC, Multithreading, Embedded Linux, REST APIs, OAuth 2.0
MLOps & DevOps: AWS SageMaker, Azure Synapse, Docker, Kubernetes, Terraform, GitHub Actions, CI/CD Pipelines
Data Engineering & Big Data: Apache Airflow, Apache Spark, PySpark, Hadoop, Hive, Kafka, Flume, Azure Data Factory, ETL Pipelines, Data Ingestion
Cloud & Infrastructure: AWS (EC2, Lambda, S3, IAM), Azure, GCP, Kubernetes, Prometheus, Grafana
Databases & Storage: PostgreSQL, MySQL, MongoDB, Redis, Elasticsearch
Frontend/Backend: Node.js, Express.js, React.js, Redux, Flask
Tools: Git, Jupyter Notebook, VS Code, OpenCV, Ansible

Work Experience

Teaching Assistant, San José State University, CA Aug. 2024 – Mar. 2025

- Mentored 60+ graduate students on AI-driven system design by integrating **C++**, **Python**, and **ROS** for robotics workflows, and enabling real-world development in **Linux-based embedded systems**. Reinforced industry-standard practices in **multithreaded programming**, fail-safe architecture, and modular deployment using simulation and validation pipelines.

AI & Data Engineer, Krasan - Microsoft, Chicago, IL, US May 2024 – Dec. 2024

- Mentored a cross-functional team while designing fault-tolerant ETL pipelines with **Airflow**, **SQL**, and **Docker**, processing over **1.2 billion records** in compliance with **DoD-style data governance** and reliability standards.
- Built embedded **forecasting models** using **TensorFlow** and **ONNX**, deployed via **Kubernetes** into **Azure Synapse**, minimizing system latency and enabling scalable integrations with existing **CI/CD pipelines** and observability tools.
- Developed a **BERT**-based NLP workflow in **PyTorch**, paired with **gRPC** and **MLflow**, to automate **real-time JD parsing**, enrich **fleet data analysis**, and support human-in-the-loop feedback for continuous improvement.

Software Engineer – Contract, Apple Inc., Hyderabad, India July 2020 – July 2023

- Designed ML pipelines in the **Banjo Portal** using **TensorFlow**, **Keras**, and **Apache Spark**, reducing false positives and enforcing **data governance**, fault tolerance, and **model accountability** for production-grade behavior models.
- Built ingestion workflows using **Apache Airflow**, **SQL**, and **Azure Data Factory**, integrating 60+ structured sources into the **Banjo Portal**, enabling performance-tuned ETL pipelines for analytics at scale.
- Improved real-time search functionality within the **Banjo Portal** using **C++**, **Boost**, and **Redis**, increasing query performance by 1.3x and enhancing enterprise data visibility for compliance and uptime reliability.
- Deployed 12 ML models with **AWS SageMaker**, leveraging **GitHub Actions** and custom **CI/CD pipelines** to ensure scalable retraining, traceability, and reproducibility across regulated environments.
- Engineered NLP chatbot workflows with **Python**, **Node.js**, and **SpaCy**, integrating with 36+ **React-Redux** components and enabling intelligent support dashboards on a **Docker**-orchestrated Kubernetes cluster.
- Provisioned secure infrastructure using **Terraform**, collaborating with DevSecOps to meet **DoD cloud compliance**, manage compute policies, and automate access control workflows for multi-environment deployment strategies.
- Mentored 8 interns in coding techniques and the entire development pipeline, earning the **Collaborative Champion Award** for strong documentation, ownership, and **cross-functional collaboration** during milestone releases.

Software Engineer, TCS Digital / Apple, Hyderabad, India July 2019 – July 2020

- Developed scalable REST APIs for the **Banjo Portal** using **Flask**, **Python**, and **SQL**, enabling secure data exchange, ingestion, and access across 15+ enterprise platforms with versioned schema tracking.
- Streamlined test workflows via **Airflow**, **Docker**, and **GitHub Actions**, automating QA environments and supporting compliance-driven **DevSecOps** initiatives with environment isolation and reproducible test automation pipelines.

Education

San Jose State University Aug. 2023 – May 2025

Masters in Computer Software Engineering – Data Science Specialization *San Jose, CA*

Coursework: Advanced Systems Programming, Applied Deep Learning, Real-Time Embedded Software, Parallel Computing with C++, Neural Computation, Sensor Fusion & Perception, High-Performance Data Pipelines, Autonomous Systems Design, Secure Edge Computing, AI Model Optimization, Robotics Software Architectures, Operating Systems for AI Applications

Relevant Projects

Please explore my GitHub for a diverse collection of projects with expertise in advanced technologies.

Hotel Rating System using Sentiment Analysis | *NLP, Python, Flask, Node.js, MongoDB* | *git-link* Aug. - Oct. 2023
* Utilized BERT LLM and NLTK NLP to analyze 40,000 hotel reviews via web scraping. Integrated **Node.js** backend and **Next.js** frontend with **MongoDB**, improving page load speed by 40%. Deployed with **CI/CD pipeline** using **Git**, reducing manual deployment time by 50%.

Twitter Data Analysis on COVID-19 using Hadoop, Flume, and Spark | *git-link* Jan. - May 2024
* Built a big data pipeline to analyze 10M+ tweets using **Apache Flume** for ingestion, **Hadoop HDFS** for storage, and **MapReduce** for querying. Processed data with **Apache Spark** and **PySpark**, benchmarking **Hive vs. Spark** on 1TB datasets. Optimized real-time search with **Elasticsearch** and **Kafka**.

Simple Cloud Infrastructure Manager (SCIM) | *git-link* Aug. - Dec. 2023
* Developed SCIM for automated cloud provisioning using **Java, C++, Terraform, and Ansible**, integrated **AWS Lambda** for serverless automation, and managed orchestration via **Kubernetes**. Secured deployments with **OAuth 2.0** and **IAM**. Enabled real-time monitoring with **Prometheus** and **Grafana**, reducing provisioning time by 50%.

Attendance Maintenance using Neural Networks, Computer Vision | *git-link* Oct. - Dec. 2019
* Developed facial scanning system with **OpenCV** for preprocessing and face-feature extraction; trained a **CNN-ANN** model achieving 92% accuracy. Implemented **spatial filtering** for clarity in low-light conditions and stored results in secured **GCP cloud** instance.

Publications and Certifications

Smart Attendance via CNN & I.P. — EJERS (IEEE Journal) May 2020
Invited speaker at ECETECH-2022, Paris, for presenting research paper.

AWS Certified Developer - Associate – Sept. 2023
Oracle Database SQL Certified Associate – Apr 2021
C++ Certified Professional Programmer (CPP) – Coursera
Machine Learning Spec. – Andrew NG – Feb. 2023
TreeHacks Hackathon – Stanford, Feb. 2024: Awarded for 'VitaVisuals' – NLP/ML health analytics project.