

Vamsi Krishna Oruganti

(562) 244-5799 | vamsikrishna.o@protectmymails.com | [LinkedIn](#) | [Google Scholar](#)

Summary

AI/ML Engineer with 4 years of experience delivering scalable machine learning and generative AI solutions using LLMs and deep learning techniques. Proficient in building end-to-end AI systems across NLP, computer vision, and predictive modeling, with hands-on deployment in cloud environments. Proven ability to align AI initiatives with business objectives and drive real-world impact.

Skills

Programming Language: Python, R Programming, SQL

IDEs: Jupyter Notebook, Google Colab

Machine Learning & Deep Learning: Linear, Logistic Regression, Decision Trees, Random Forests, Naive Bayes, K-Means, CNN, RNN, ANN, LSTM, DBSCAN, Principal Component Analysis (PCA), Autoencoders

Deep Learning Framework: Tensorflow, Keras, PyTorch

AI Technology: LangChain, LLaMA, Mistral, Large Language Model (LLM), LlamaIndex, GANs, Retrieval-Augmented Generation (RAG)

NLP Models: ROBERTa, Claude, BERT, GPT-4, Named Entity Recognition (NER)

Cloud: AWS (SageMaker, Lambda, CloudWatch, CloudFormation), **Azure** (Azure Kubernetes Service, Azure Blob Storage)

Libraries: PySpark, OpenCV, Scikit-learn, Pandas, NumPy, SciPy, Matplotlib, NLTK, Spacy, Seaborn, XGBoost

Visualizations and Version Control Tools: Tableau, Power BI, Git, Bitbucket

Database: SQL Server, MySQL, PostgreSQL, MongoDB, Firebase

Education

Master of Science in Computer Science

California State University, Long Beach, CA, USA

2024

GPA: 3.5/4.0

Bachelor of Technology in Computer Science and Engineering

SRM University, AP, India

2022

GPA: 8.3/10.0

Work Experience

PNC Financial Services, USA

Apr 2025 – Current

AI/Machine Learning Engineer

- Led the development of a deep learning-based fraud detection system using Convolutional Neural Networks and DBSCAN to combat transactional anomalies, increasing fraud identification accuracy by 19%, thereby enhancing digital banking security and customer trust.
- Built a Retrieval-Augmented Generation (RAG) system utilizing LangChain and LlamaIndex to streamline financial document retrieval for customer service agents, cutting average resolution time by 45% and improving first-contact resolution in client communications.
- Created Tableau dashboards visualizing AI-driven insights on fraud risk, customer behavior, and loan segmentation, empowering senior leadership to make proactive, data-backed decisions in retail and commercial banking operations.
- Enhanced forecasting and sentiment analysis models by implementing advanced preprocessing, Focal Loss to address class imbalance, and dynamic learning rate schedulers, resulting in significantly improved model accuracy and stability in financial datasets.
- Integrated Mistral LLMs into a financial advisor chatbot, delivering real-time investment insights that enhanced advisor productivity and maintained a 90% satisfaction rate from internal users, reinforcing commitment to digital transformation in wealth management.

Kings Hawaiian Holding, Inc., Gardena, California

Aug 2024 – Mar 2025

Machine Learning Engineer

- Developed a financial forecasting tool using Python and DataRobot to predict cash flows, achieving ~90% accuracy and helping the finance team improve budgeting for seasonal inventory demands.
- Cleaned and transformed raw sales, inventory, and expense data from multiple departments to prepare structured datasets for modeling, reducing data preparation time by 30%.
- Built and tested regression and time series models (ARIMA) to analyze historical trends in product demand, enabling more accurate forecasting during promotional campaigns and holidays.
- Designed and trained Convolutional Neural Networks (CNNs) to analyze in-store product images for shelf placement compliance, improving retail audit accuracy by 35%.
- Applied Principal Component Analysis (PCA) to reduce dimensionality of operational datasets, resulting in a 40% reduction in model training time without sacrificing accuracy.
- Integrated AWS S3 and SageMaker for scalable model training and deployment pipelines, improving reliability and reducing manual intervention during retraining cycles.

Allstate, USA

Sep 2023 – May 2024

Data Scientist

- Developed dynamic Tableau dashboards integrating predictive models with policyholder spend and claims data, enabling real-time tracking of cost overruns and policy leakage, which reduced manual reporting effort and improved financial transparency.
- Built predictive analytics models using Decision Trees, Random Forests, Naive Bayes, and SVM to forecast insurance product demand and regional quote activity, increasing forecasting accuracy by 18% and optimizing resource planning for underwriting teams.
- Deployed NLTK-based sentiment analysis and text summarization pipelines on unstructured customer feedback from claims and support channels, uncovering pain points and driving a 15% increase in customer satisfaction through targeted service enhancements.
- Leveraged Scikit-learn to rapidly prototype, train, and validate multiple risk scoring models, cutting model development time while ensuring compliance with internal model governance standards.
- Engineered a claims severity prediction model by combining structured adjuster reports and NLP-derived insights from claimant statements, improving triage efficiency and reducing claim settlement time.
- Collaborated with actuarial and fraud analytics teams to analyze anomalies in premium payments and claim ratios using clustering algorithms, helping uncover previously undetected fraud patterns worth over \$500K in potential exposure.

Kpit Technologies, India

Jun 2020 – Jul 2022

Data Scientist / Jr. Machine Learning Engineer

- Designed vehicle usage-based segmentation models using K-Means and DBSCAN on telematics data, identifying driver behavior clusters that enabled predictive maintenance campaigns, increasing campaign ROI by 15% for a leading OEM client.
- Optimized model pipelines for connected vehicle analytics by applying PCA for dimensionality reduction and engineering domain-specific features in Python, reducing training time by 35% without compromising accuracy in mobility pattern predictions.
- Created deep learning models with CNNs and RNNs in TensorFlow to classify automotive service logs and forecast vehicle part failures using time-series data, achieving 92% accuracy and enabling automated diagnostics in ADAS systems.
- Developed and deployed demand forecasting models using ARIMA and XGBoost on historical vehicle part usage data, improving inventory planning accuracy by 21% and minimizing excess stock across dealer networks.
- Created Power BI dashboards to visualize predictive model insights across automotive lifecycles, maintenance, warranty, and diagnostics, helping stakeholders make faster, data-driven operational decisions.
- Implemented scalable ML pipelines on Azure Kubernetes Service (AKS) integrated with Azure Blob Storage for real-time diagnostics model deployment, ensuring low-latency inference and reducing cloud costs by 17%.
- Performed extensive exploratory data analysis (EDA) on automotive datasets using Python (Pandas, NumPy, Seaborn), uncovering patterns like fuel inefficiency triggers and outliers in vehicle telemetry, which improved feature selection and cut preprocessing time.

Projects

Speech Emotion Recognition | Python, MLP Classifier, NLTK

- Designed and implemented a Speech Emotion Recognition system using Python and Librosa to extract MFCCs and other key audio features from English audio data, enabling accurate emotion classification across categories like happiness, sadness, anger, and neutrality.
- Trained an MLP (Multi-Layer Perceptron) classifier on the engineered feature set and applied NLTK for text pre-processing of transcriptions, achieving a validated accuracy of 72.4%, effectively enhancing emotion-aware applications such as virtual assistants and voice-based sentiment systems.

Skin Cancer Detection | Python, CNN, PyTorch, TensorFlow

- Developed a CNN-based skin lesion classification system using the ISIC2018 dataset, enhancing diagnostic accuracy by applying ESRGAN for high-resolution image preprocessing and extracting robust visual patterns indicative of malignancy.
- Implemented transfer learning with ResNet50, InceptionV3, and Inception ResNet architectures using PyTorch and TensorFlow, optimizing model performance through advanced data augmentation, normalization, and input resizing techniques for improved generalization on clinical image data.

Internal Knowledge Chatbot | Python, Lang Chain, OpenAI GPT-4, FAISS, Streamlit

- Built a data-driven chatbot platform to extract and index knowledge from internal documents (PDF/DOCX), enabling efficient information retrieval with a Retrieval-Augmented Generation (RAG) pipeline.
- Engineered data ingestion, document chunking, semantic indexing using FAISS, and integrated GPT-4 for context-aware data queries. Enabled real-time data access and retrieval over knowledge bases, improving internal data search.

Achievements

a) Published Research papers

1. <https://trid.trb.org/View/2401722>
2. <https://fersc.utk.edu/wp-content/uploads/sites/8/2024/09/FERSC-2023-Project7-1.pdf>

b) Presented a research topic at FERSC, 2024, at Texas A&M University

c) Worked as a Teaching Associate for a MATLAB course and taught undergrad students