

ANTHONY KOMMAREDDY

📞 669-799-7863 | ✉ anthony.sandesh.k@gmail.com | 🌐 anthony.sandesh.com | 🔗 linkedin | 🐙 github | 📍 San Jose, CA 95126

Summary

Experienced AI Engineer and Research Assistant with a strong background in machine learning, large language models, and data analysis. Skilled in developing and deploying AI solutions that drive business growth and operational efficiency.

Experience

AI Engineer: BhizChat INC, San Jose, California

Oct 2024 – Present

- Engineered and deployed LLM models and NLP solutions via Docker and Jenkins **CI/CD pipelines**, optimizing workflows, increasing conversion rates by 30%, and reducing cart abandonment by 25% for 3,000+ users.
- Executed **GPT-4**-powered product recommendation systems, incorporating real-time data analysis with tools such as Pandas and TensorFlow, driving a 40% improvement in product discovery and satisfaction.
- Designed an analytics dashboard using Streamlit and Python, incorporating A/B testing frameworks to optimize chat interactions, boosting average order values by 20% for 500+ businesses.

Research Assistant: SJSU Research Foundation, San Jose, California

Jul 2024 – Present

- Created an interactive intrusion detection dashboard leveraging Random Forest classifiers and integrating **SHAP** and **LIME** for explainable AI insights on the **KDD99** dataset.
- Orchestrated the integration of **GPT-3.5** Turbo conversational agents, translating threat detection into clear, actionable insights, streamlining analyst workflows.
- Evaluated system accuracy using **CISM** best practices, employing precision-recall metrics and Python libraries like Scikit-learn to deliver actionable insights with high precision.

Research Scientist Intern: Ericsson, Stockholm, Sweden

Jan 2023 – Jul 2023

- Optimized error detection in log files by employing **SVM**, **multinomial Naive Bayes**, and **LSTM** models, utilizing TensorFlow and Scikit-learn, streamlining system integration and verification processes.
- Conceptualized and devised verdict generation algorithms leveraging **Luhn**, **LSA**, and **LexRank**, benchmarking results against **GPT-3.0** to automate log classification, reducing manual efforts by 50%.
- Constructed an anomaly detection pipeline for large-scale log datasets using **Regex**, TensorFlow, and LLMs, enhancing operational efficiency by 30% and accuracy rates significantly.

Machine Learning Intern: Intone INC, Hyderabad, India

Jan 2023 – Jul 2023

- Improved risk classification accuracy by 25% through **TensorFlow** and Scikit-learn, leveraging big data frameworks like Apache Spark and Hadoop for efficient model training and deployment.
- Developed and deployed an OCR-based cheque and invoice reader utilizing OpenCV and **Tesseract**, accelerating data processing by 40% and automating anomaly detection using NLP models.
- Designed and implemented a facial recognition attendance system leveraging **YOLOv8** and **OpenCV**, streamlining attendance management workflows and boosting operational efficiency by 20%.

Projects

Personal Finance Analysis | LLMs, LLaMA, Vector Database, ChromaDB, Streamlit

Jul 2024 - Aug 2024

- Engineered a tool using **Local LLMs** with 90% accuracy for transaction classification.
- Integrated **PDF loaders** and **vector databases** for scalable analysis and predictive financial insights.

Smishing Attack Detection in Messages | — RoBERTa, Transformers, Streamlit

Jan 2024 - Mar 2024

- Engineered a detection model, achieving 95% accuracy and reducing attack risks by 60% through advanced filtering techniques.
- Built an **interactive UI** using Streamlit, providing 40% faster result visualization for enhanced user experience.

Education

San Jose State University

Jan 2024 – Dec 2025

Master of Science in Artificial Intelligence

San Jose, CA

Blekinge Institute of Technology

Aug 2018 – May 2022

Master of Science in Computer Science and Engineering

Karlskrona, India

Course Work: Recommender Systems, Data Mining, Artificial Intelligence, Deep Learning, Machine Learning Cloud Computing, Object-oriented programming, Data Structures & Algorithms, Database Management Systems, Web UI Programming

Skills

Languages: Python, C, C++, HTML, CSS, Java, JavaScript, TypeScript, SQL, C++, Java

Frameworks: PyTorch, Tensorflow, Flask, Keras, Pandas, OpenCV, RESTful APIs, scikit learn, Node.js, React.js

Technologies: Linux, MongoDB, Postgres, MySQL, AWS, Git, LLMs, GenAI, AWS, GCP, Azure

AI / ML: TensorFlow, PyTorch, Keras, Transformers, LangChain, MLOps, LLMops, CV, NLP, Matplotlib, Streamlit, Tableau

Developer Tools: Github, Docker, Kubernetes, GPUs, Selenium, Agile, Jira, Postman

Publication

[1] Kommareddy, Anthony. “**System Integration and Verification Verdict Automation Using Machine Learning**” Presented at BTH, 2023. Conducted under **Dr. Prashant Goswami** and Ericsson advisors **Jerker Zetterlund** and **Sike Huang**, this research explores reducing system integration time for the Cloud RAN project using machine learning and NLP and automated verification process that improved error detection and developed a web application to optimize workflows.