Eswar Sai Korrapati

551-344-5356 | eksai0726@gmail.com | LinkedIn | Github | Portfolio

Professional Summary

AI/ML Engineer with more than three years of experience building production-ready machine learning and agentic AI systems. Skilled in designing scalable pipelines on AWS with SageMaker, MLflow, and Spark. Strong background in API design, automation, and MLOps practices that keep workflows reproducible and reliable. Experienced in model explainability with SHAP and LIME and in tuning models for fast, real-time predictions in production.

Experience

Software Engineer – AI Oct 2025 – Present
eSentire *Pleasanton, CA*

- Built a supervisor-based agent workflow that generates structured JSON for dynamic security dashboards.
- Designed JSON outputs with data endpoints, UI components, and layout details for dashboard rendering.
- Used RAG to pull schema details and UI component metadata for each client.
- Implemented strict prompt-engineering rules to enforce structure and keep agent outputs stable.
- Deployed agent workflows on AWS Lambda and monitored performance through CloudWatch.
- Collaborated with platform teams to improve deployment pipelines using Kubernetes, increasing dashboard reliability and system uptime.

Software Engineer – AI Aug 2024 – Sep 2025
Comcast *San Jose, CA*

- Developed agentic AI workflows with LangChain and LangGraph using Claude Sonnet 4 to improve system efficiency and automate support tasks.
- Added persistent state to multi-agent flows, reducing repeated interactions by more than 30%.
- Built supervisor–worker flows that improved automated response accuracy and reduced escalations.
- Integrated tools through MCP so agents could run internal lookups and troubleshooting steps, increasing operational coverage.
- Deployed workflows on AWS Lambda, monitored them through CloudWatch, and reduced rollout time and operational cost.
- Worked with Agile teams to streamline CI/CD cycles and improved deployment efficiency using Kubernetes.

Software Engineer May 2020 – Jul 2022
Dentsu *Hyderabad, India*

- Built Flask APIs supporting marketing analytics, improving data processing speed and decision-making.
- Developed XGBoost models for ad-campaign forecasting with more stable and accurate predictions.
- Used PySpark for data cleaning, joins, and feature preparation across large datasets.
- Versioned data and models with DVC to ensure consistency and reproducibility across ML experiments.
- Built Airflow pipelines for ETL and retraining, reducing manual restarts and improving reliability.
- Worked with Agile teams to deploy SageMaker endpoints and Docker-based EC2 services, improving scalability and uptime.

Education

Montclair State University Montclair, NJ
Master's in Computer Science, GPA: 3.65 *May 2024*

Projects

RAG-based PDF Summarizer | Link
Next Word Prediction with BERT Transformer | Link
Machine Transliteration | Link
AI Meeting Preparation Agent | Link

Technical Skills

Programming Languages: Python, C++, SQL, JavaScript
AI/ML & NLP: TensorFlow, PyTorch, Scikit-learn, Keras, Transformers, Hugging Face, LangChain, LangGraph, n8n
MLOps & Deployment: MLflow, SageMaker, DVC, Airflow, Docker, Kubernetes, CI/CD, Jenkins
Explainability & Evaluation: SHAP, LIME, AUC, Precision, Recall, F1 Score, Confusion Matrix
Cloud & DevOps: AWS (S3, EC2, Lambda, SageMaker, CloudWatch, Bedrock), GitHub Actions
Data & Processing: Apache Spark, Pandas, NumPy, MongoDB, MySQL, Tableau
LLMs & APIs: OpenAI API, Prompt Engineering, FastAPI, REST APIs, RAG, MCP
Tools & Collaboration: Git, Jupyter, Streamlit, Agile, IntelliJ, VS Code, Cursor
Core CS: Data Structures and Algorithms