Swarnalatha Muppavarapu

Frisco, TX · swarna.mp01@gmail.com · +1 (510) 427-2514 ·
https://www.linkedin.com/in/swarnalatha-muppavarapu/

Education

Mamata College — India
Master's Degree in Science — 2008 - 2011

Experience

Gainwell Technologies — Irving, TX
AI/ML Engineer — Jan 2023 - present

- Developed AI-powered chatbots using Generative AI and Large Language Models to deliver personalized, context-aware responses, improving user engagement and satisfaction.
- Applied NLP techniques with transformer-based models (BERT, GPT) for natural language understanding and generation, enhancing chatbot accuracy and contextual relevance.
- Designed and implemented Agentic AI workflows using LangGraph and LangChain, enabling autonomous decision-making, dynamic tool orchestration, and adaptive multi-step reasoning.
- Built modular and stateful agent architectures to manage evolving user contexts, long-term objectives, and strategy switching for improved task completion.
- Integrated Agentic AI with external APIs and knowledge sources, enabling real-time retrieval, reasoning, and contextual grounding for enhanced adaptability.
- Implemented Retrieval-Augmented Generation (RAG) pipelines using vector databases (Pinecone, Milvus) to enhance response relevance with external knowledge.
- Developed evaluation frameworks using RAGAS and BERTScore to assess retrieval quality, factual consistency, and semantic similarity of chatbot responses.
- Fine-tuned transformer models (BERT for classification, BART for summarization) and leveraged Hugging Face Transformers for domain-specific NLP tasks.
- Deployed scalable AI solutions on AWS (Bedrock, EKS) with RESTful APIs (FastAPI) for real-time inference and efficient model serving.
- Implemented LLMOps pipelines using Docker and CI/CD workflows for versioning, deployment, and monitoring of large language models.
- Collaborated with cross-functional teams to integrate domain knowledge into AI models, aligning technical solutions with business requirements.
- Optimized model performance and inference latency through efficient embedding management and vector search strategies, improving real-time responsiveness.
- Enhanced chatbot reasoning by integrating multi-step tool orchestration and autonomous decision-making workflows for complex user queries.
- Leveraged data-driven insights to continuously refine model outputs and maintain high-quality, contextually accurate responses.
- Developed scalable, production-ready AI solutions with monitoring and logging for reliability and maintainability in enterprise environments.
- Implemented conversational analytics to monitor user interactions, identify improvement areas, and enhance overall user experience.
- Integrated security and compliance best practices in AI workflows to ensure data privacy and regulatory adherence.
- Explored and applied emerging AI techniques to continuously innovate chatbot capabilities and stay ahead in generative AI trends.

UHG — Newark,CA
Machine Learning Engineer — July 2019 - Dec 2022

- Implemented ML pipelines for clustering and segmentation of large-scale datasets, leveraging AWS for efficient querying and integration with ETL workflows.
- Built scalable and automated ML pipelines for data preprocessing, feature engineering, and model training using pandas, scikit-learn, and NumPy, ensuring high-quality data transformation and modeling.

- Applied MLOps best practices using MLflow, DVC, and Databricks to manage experiments, version control, and dataset tracking, enhancing collaboration and reproducibility across teams.
- Integrated Docker, Kubernetes, and TorchServe for deploying models on AWS SageMaker, ensuring scalable and reliable serving infrastructure for production environments.
- Designed CI/CD pipelines for automated testing, model deployment, and updates, streamlining machine learning workflows and reducing operational overhead.
- Developed scalable data pipelines using AWS services such as Lambda and S3, enabling real-time data ingestion and seamless integration with machine learning models.
- Implemented model monitoring and alerting systems on AWS (CloudWatch, SageMaker Model Monitor), tracking key performance metrics to ensure accuracy, stability, and minimal downtime
- Collaborated with cross-functional engineering and data teams to integrate ML models into core systems, improving prediction accuracy, system efficiency, and business impact.

AMEX Palo Alto, CA
Software Engineer May 2018 - June 2019

- Designed, developed, and maintained scalable software applications using modern programming languages.
- Implemented RESTful APIs and integrated third-party services to enhance system functionality.
- Developed automated unit, integration, and functional tests to ensure software quality.
- Collaborated with cross-functional teams to gather requirements and translate them into technical solutions.
- Optimized application performance and improved system reliability through code reviews and refactoring.
- Participated in Agile/Scrum development processes, including sprint planning and daily stand-ups.
- Analyzed and debugged complex software issues, providing timely resolutions and improvements.
- Conducted code documentation, maintained design specifications, and shared best practices across the team.
- Assisted in deploying applications to cloud and on-premises environments while ensuring security and compliance standards.

CIGNA Health Insurance Southfield, MI
Software Engineer Jan 2017 - April 2018

- Designed, developed, and maintained enterprise web applications to support healthcare business operations.
- Implemented application features using Java, integrated with SOAP/RESTful APIs, and validated request/response structures in XML/JSON.
- Automated unit and integration testing using JUnit and contributed to CI/CD pipeline improvements to ensure code reliability.
- Wrote SQL queries, stored procedures, and triggers to validate data integrity and enforce business logic in relational databases.
- Collaborated with cross-functional teams in an Agile/Scrum environment for requirement analysis, sprint planning, and defect resolution.
- Debugged and deployed applications on UNIX-based platforms, optimizing system performance and availability.
- Developed and executed test scenarios and regression test suites to validate system functionality and reduce production defects.
- Partnered with QA teams by creating test data, executing validation checks, and assisting in defect triage to accelerate release cycles.

Skills

Programming Languages:	Python, Java
ML Frameworks/Libraries:	Pytorch, Huggingface, MLFlow
Cloud Platforms:	AWS Bedrock, AWS Sagemaker
Data Processing and Analysis:	Numpy, Pandas, Spark, Databricks
Natural Language Processing (NLP):	NLP, Pytorch, Huggingface
Web Development and Deployment:	Streamlit, FASTAPI, Docker, Kubernetes
QA Tool And Tech:	Selenium RC, Unit, SOAPUI, SQL, QTP, QC, UNIX.
Other Tools and Technologies:	LangChain/LlamaIndex, RAG, Gen AI, LLM, Vector DB