

Bhargavi Kadambari

TX, USA | 469-351-9574 | bhargavikadambari32@gmail.com | [LinkedIn](#) | <https://github.com/bhargavi32>

SUMMARY:

AI/ML Engineer with 4+ years of experience in cloud data engineering and advanced machine learning solutions. Strong expertise in LLMs, NLP, GenAI, deep learning, and RAG architectures, with hands-on work in LangChain, vector databases, knowledge graphs, and MLOps. Skilled at building real-time AI systems for fraud detection, interview agents, and predictive analytics using AWS & GCP.

SKILLS:

- **AI/ML Core:** LLMs, NLP, Computer Vision, GenAI, RAG Architecture, AI Agents, Vector Embeddings, LLM Fine-Tuning, Causal Inference, Bayesian Methods, Statistical Analysis, A/B Testing, ML Algorithms, Time Series Forecasting, Recommendation Systems.
- **Frameworks & Tools:** PyTorch, TensorFlow, Keras, Scikit-Learn, HuggingFace Transformers, LangChain, LLAMA, XGBoost, Model Context Protocol (MCP), OpenAI, DeepSeek, Uma.
- **Deep Learning:** ANN, CNN, RNN, LSTM, Transformers, BERT, LoRA, QLoRA, Multi-Head Attention Mechanisms, Neural Network Optimization, Few-Shot Learning, GANs.
- **Data Science & Analytics:** EDA, Data Cleaning, Preprocessing, Feature Engineering, Statistics, Matplotlib, Seaborn, Excel Charts, Tableau, Looker Studio.
- **GenAI & NLP Advanced:** LLM Deployment, Prompt Engineering, Chain-of-Thought, Constitutional AI, Semantic Caching, Knowledge Graphs.
- **MLOps & Engineering:** End-to-End ML Pipelines, Model Training, Testing & Deployment, CI/CD, Docker, Kubernetes, Containerization, Model Monitoring, Model Governance, Git, GitHub, GitHub Copilot, Bitbucket, Distributed Computing.
- **Cloud & Infrastructure:** AWS (S3, EC2, EKS, ECS, ECR, Lambda, Bedrock, SageMaker, EMR, Athena, Batch, Step Functions, Redis), Azure (AML, AKS), Snowflake, Microservices.
- **Data Engineering:** ETL Pipelines, Big Data Processing, Hadoop, Spark, Kafka, Apache Airflow, Real-Time Systems, Data Governance.
- **Databases:** MySQL, NoSQL, Vector Databases (Pinecone, ChromaDB).
- **Programming:** Python, SQL, PySpark, R, OOP Concepts, Numpy, Pandas.
- **Operating Systems:** Linux, UNIX, Windows, MacOS.
- **Developer & Productivity Tools:** Cursor, Windsurf, IntelliJ, Visual Studio, GitHub Copilot.

PROFESSIONAL EXPERIENCE:

Invent Artificial LLC – AI/ML Engineer Intern

Jan 2025 – Present

- Built RAG pipeline with Pinecone/ChromaDB, achieving 92% contextual accuracy in Q&A.
- Integrated MCP tools & knowledge graphs, enabling domain-adaptive interview agents.
- Developed multi-modal STT + LLM pipeline, boosting engagement by 40%, and optimized inference latency by 60% via semantic caching.
- Fine-tuned LLMs (OpenAI, DeepSeek, and Uma) with LoRA/QLoRA and applied constitutional AI guardrails for ethical alignment.
- Deployed Dockerized microservices on AWS (EKS, ECS, Lambda, SageMaker) with CI/CD for scalable delivery.
- Designed an AI Interview Agent with TTS + real-time evaluation models (clarity, correctness, confidence).

Data Engineer – GCP| Wipro Technologies

2021 – 2023

- Designed and implemented ETL pipelines using Dataflow (Apache Beam) and Dataproc (PySpark) for large-scale batch/streaming workloads.

- Migrated legacy systems to BigQuery, applying clustering/partitioning to cut query latency by 30% and improve cost efficiency.
- Built real-time streaming pipelines with Pub/Sub + Dataflow, enabling anomaly detection and predictive analytics with near-zero latency.
- Automated workflow orchestration with Cloud Composer (Airflow) and deployed data quality checks with GCP Data Catalog, ensuring ML-ready datasets and governance compliance.

Data Engineer – AWS | Wipro Technologies |

2021 – 2022

- Architected a cloud-based data lake on Amazon S3 and automated ETL workflows with Glue + Step Functions, reducing processing times by 45%.
- Built real-time ingestion pipelines with Kafka + AWS Lambda, improving reporting latency by 40%.
- Integrated Amazon Redshift with optimized transformations and incremental loading, boosting query performance by 35%.
- Implemented Airflow orchestration and pipeline monitoring, ensuring reliability, compliance, and scalable analytics delivery.

Application Engineer – TCS iON

2020 – 2021

Responsibilities:

- Developed and deployed containerized Java applications using Docker, ensuring portability and consistency across development, testing, and production environments.
- Configured and maintained Docker Compose for orchestrating multi-container applications, streamlining integration of databases, APIs, and services.
- Automated build, testing, and deployment pipelines using CI/CD workflows, reducing manual effort and cutting deployment time by 40%.
- Integrated Git-based version control and container registries to manage and distribute container images.
- Conducted performance optimization and container lifecycle management, improving system efficiency and resource utilization.

ACADEMIC PROJECTS

FinGuard – AI Fraud Detection Platform:

- Designed and deployed a fraud detection system using ensemble ML models (Logistic Regression, Random Forest, Neural Networks), achieving 99% detection accuracy on high-volume transaction data
- Implemented Explainable AI (XAI) techniques with SHAP and LIME for transparent fraud scoring and regulatory compliance

NYC 311 Data Pipeline & Forecasting:

- Designed an end-to-end cloud data pipeline on Google Cloud Platform (BigQuery, Dataflow, Dataproc, Cloud Composer) to process millions of NYC 311 service requests
- Automated ETL workflows using PySpark on Dataproc, handling structured and unstructured datasets with data validation and governance checks

CERTIFICATIONS:

- Microsoft Certified: Azure AI Fundamentals (AI-900)
- Google Cloud Certified Professional Data Engineer
- Oracle Database SQL Certified Associate (1Z0-071)

EDUCATION:

- | | |
|--|-----------------------------|
| • University Of North Texas
Masters in Data Science | Jan 2024 - Present |
| • Jawaharlal Nehru Technological University
Bachelors in Electronics & Communication | July 2017 - Jun 2021 |