Likhitha Bompelly

New York, NY 13417

📞 (984) 281-9343 ✉ likhithabompelly276@gmail.com 🔗 [LinkedIn](#)

PROFESSIONAL SUMMARY

Generative AI Engineer with 5+ years of experience in machine learning, deep learning, and large language models. Skilled in designing, fine-tuning, and deploying generative AI applications across text, image, and NLP domains. Adept at integrating LLMs into enterprise workflows, building scalable ML pipelines, and ensuring ethical, explainable AI. Seeking to leverage expertise in AI engineering to deliver transformative business solutions.

TECHNICAL SKILLS

- **Core AI/ML**: Deep Learning (CNN, RNN, Transformers), Generative Models (GANs, VAEs, Diffusion), NLP, Reinforcement Learning, Transfer Learning
- **Generative AI Frameworks**: OpenAI API (GPT-4, DALL·E, Codex), Hugging Face Transformers, LangChain, LLaMA/Mistral/Falcon, LoRA/PEFT, Stable Diffusion
- **Tools & Platforms**: Jupyter, Colab, MLflow, Weights & Biases, Streamlit, Gradio, FastAPI, Flask
- **Programming**: Python (NumPy, Pandas, PyTorch, TensorFlow, Keras), SQL, Bash, JavaScript, C++
- **Cloud & MLOps**: AWS SageMaker, Azure ML, Google Vertex AI, Bedrock, Docker, Kubernetes, GitHub Actions, Jenkins
- **Data Engineering**: ETL Pipelines, Data Lakes, Tokenization, Feature Engineering, Model Optimization
- **Other Skills**: RAG (Retrieval-Augmented Generation), Vector Databases (Pinecone, FAISS, Weaviate), Knowledge Graphs, SHAP, LIME, Responsible AI

PROFESSIONAL EXPERIENCE

M&T Bank   **Jan 2025 – Present**
Role: Generative AI Engineer   New York, USA

Description: At M&T Bank, Delivered enterprise-scale generative AI solutions, integrating LLMs into financial applications for risk analysis and customer engagement.

Responsibilities:

- ➢ Designed, fine-tuned, and deployed GPT-based applications for banking operations.
- ➢ Built prompt engineering frameworks to standardize LLM outputs, reducing hallucinations.
- ➢ Developed data pipelines for preprocessing and cleansing structured/unstructured data.
- ➢ Integrated LLM APIs with FastAPI microservices, deployed using Docker & CI/CD pipelines.
- ➢ Built predictive models using Random Forest, SVM, and deep learning architectures.
- ➢ Automated ML workflows on AWS SageMaker with GitLab integration for deployment.
- ➢ Created real-time monitoring dashboards with Tableau and Datadog.
- ➢ Conducted exploratory data analysis (EDA) and visualization for model insights.
- ➢ Collaborated with cross-functional teams to align AI solutions with business needs.
- ➢ Mentored junior engineers on prompt engineering and MLOps best practices.
- ➢ Applied bias mitigation and compliance checks for responsible AI use.
- ➢ Delivered technical documentation and playbooks for model governance.
- ➢ **Environment**: Python, NumPy, Pandas, Scikit-learn, TensorFlow, PyTorch, AWS SageMaker, GitLab, Tableau, FastAPI, Docker, CI/CD Pipelines, LLM APIs

New York Life Insurance Company   **Feb 2024 – Dec 2024**
Role: Machine learning Engineer   New York, USA

Description: At New York Life, Developed and deployed predictive models and Enhanced customer service systems

with AI-driven insights and generative AI applications.

Responsibilities:

- Built NLP and LLM pipelines for insurance workflows using Azure ML and Vertex AI.
- Standardized prompt templates for GenAI applications, ensuring consistency and compliance.
- Designed and deployed ML models for policy recommendation and risk forecasting.
- Implemented model monitoring for drift detection and versioning with CI/CD pipelines.
- Collaborated with product owners to translate requirements into AI use cases.
- Applied predictive analytics for customer churn and sales forecasting (90%+ accuracy).
- Built SQL/ETL pipelines for large-scale data ingestion and preprocessing.
- Delivered Tableau dashboards and visualizations for business stakeholders.
- Implemented explainability methods (SHAP, LIME) to improve trust in models.
- Ensured compliance with bias mitigation and data governance policies.
- Automated reporting processes with Python and R scripts.
- Supported model deployment in hybrid cloud environments.
- **Environment**: NLP, Azure ML, Vertex AI, Python, R, SQL, Tableau, CI/CD, LLM Pipelines, Responsible AI Tools

Broadridge Financial Solutions   **July 2022 – Dec 2022**

 Role: Data Scientist   Mumbai, India

Description: Developed the data analysis model according to business scenarios, Evaluate the method and technical solution of data analytics projects

Responsibilities:

- Applied supervised, unsupervised, and reinforcement learning to business forecasting.
- Developed predictive models for sales and customer segmentation.
- Built Azure LUIS NLP models for entity recognition and chatbots.
- Created data pipelines in Hadoop, Spark, and Hive for large-scale analytics.
- Designed dashboards in Tableau for senior leadership reporting.
- Implemented anomaly detection models using Scikit-learn.
- Built regression and decision-tree models for business outcome prediction.
- Collaborated with stakeholders to translate business needs into AI solutions.
- Delivered technical reports and knowledge transfer sessions.
- Developed chatbot solutions with Azure Bot Framework.
- **Environment**: Python, R, SAS, SQL, Azure LUIS, Tableau, Hadoop, Spark, Hive, Kafka, Scikit-learn

Amazon.com, Inc.   **Jan 2020 - Mar 2022**

Role: Data Scientist   Mumbai, India

Description: Developed fraud detection and operational applications within the e-commerce ecosystemI delivered some of most strategic technical projects, deliver large scalable products, design new products as Confidential at the cutting edge of distributed technology, using AI tools to deliver results.

 Responsibilities:

- Built fraud detection and credit risk analytics systems using Apache Spark & Ignite.
- Developed LSTM and GRU models for text classification and sentiment analysis.
- Deployed ML models as web services via AWS (EC2, S3, DynamoDB, Glue).
- Created predictive analytics pipelines for customer behavior forecasting.
- Designed real-time fraud prevention solutions for high-volume transactions.
- Built deep learning models with Keras and TensorFlow for classification tasks.
- Partnered with product managers to align AI solutions with customer needs.
- Conducted anomaly detection and customer trend analysis using ML.
- Mentored junior analysts on Spark-based data processing.
- Authored technical documentation for AI toolkits and workflows.
- **Environment**: AWS (S3, EC2, DynamoDB), Apache Spark, Apache Ignite, Python, SQL, TensorFlow, Keras, ML Toolkit

EDUCATION

SUNY Polytechnic Institute – New York, USA   **Jan 2023 - Dec 2024**
Master of Science, Computer Science
Coursework : Advanced Machine Learning, Deep Learning, Natural Language Processing, Data Mining, Big Data Analytics, Distributed Systems, Cloud Computing
Jawaharlal Nehru Technological University - Hyderabad, India   **Aug 2018 – July 2022**
Bachelor's in computer science & engineering
Coursework: Algorithms, Database Systems, Artificial Intelligence, Operating Systems, Data Structures.

CERTIFICATIONS & AWARDS

- Microsoft Certified: Azure AI Engineer Associate
- AWS Certified Machine Learning – Specialty
- Google Cloud Professional Machine Learning Engineer
- Databricks Generative AI Fundamentals