Prashanth Madishetti
+1 302-444-1912
prashanthmadishetti8@gmail.com
Gen AI Engineer

PROFESSIONAL SUMMARY:

- 9+ years of experience in **Generative AI**, **AI/ML**, and **Data Engineering**, delivering scalable AI solutions that drive innovation and enterprise efficiency.
- Skilled in working with **LLMs** (**BERT, T5, LaMDA, BLOOM**) and frameworks like **LangChain** for NLP/NLG tasks, including summarization, conversational AI, and content generation.
- Proficient in **Generative AI models**, including **GANs** and **Diffusion Models**, and implementing semantic search using **vector databases** like **Pinecone**.
- Strong expertise in deploying **AI workflows** with **GCP Vertex AI**, **Kubernetes**, and **CI/CD pipelines** for production-grade scalability and reliability.
- Hands-on experience with **Text-to-Speech** (**TTS**), **voice cloning**, and **speech recognition** using **PyTorch**, **TensorFlow**, and toolkits such as **Librosa** and **ESPnet** for building multilingual, real-time conversational AI applications.
- Knowledgeable in advanced machine learning techniques such as **Convolutional Neural Networks (CNNs)**, **Recurrent Neural Networks (RNNs)** and deep learning with PyTorch and TensorFlow for predictive analytics, anomaly detection, and unstructured data processing.
- Experienced in **Explainable AI techniques** (SHAP, LIME) for improving model transparency, interpretability, and compliance in enterprise AI systems.
- Skilled in building **predictive models** using **XGBoost**, **SVM**, and **Random Forest**, and **deep learning** with PyTorch and TensorFlow for structured and unstructured data.
- Extensive **Python expertise**, leveraging Pandas, NumPy, SciPy, Scikit-learn, Matplotlib, and Seaborn for advanced data analytics and visualization.
- Experienced in **Python-based AI/ML pipelines**, integrating **PyTorch**, **TensorFlow**, and Hugging Face Transformers for building and deploying LLMs and Generative AI solutions.
- Proficient in **Python automation**, using scripts to streamline data preprocessing, model training, evaluation, and deployment in GCP Vertex AI, **AWS SageMaker**, and **Azure ML**.
- Hands-on experience in automating ML workflows with **Vertex AI Pipelines** and MLOps practices for scalable training, deployment, and monitoring.
- Experience in designing monitoring strategies for deployed ML models using **Vertex AI** Model Monitoring and Cloud Logging.
- Extensive experience in ETL and data engineering, designing pipelines with **Apache NiFi**, **Airflow**, **Azure Data Factory (ADF)**, and **Azure Databricks**.
- Strong expertise in big data ecosystems, including **Hadoop, Spark, Hive,** and real-time streaming pipelines (**Kafka, Apache Beam**) for high-throughput data processing.
- Proficient in **SQL/T-SQL development** and data warehouse schema design, optimizing analytics and reporting workflows.
- Strong Experience in EDA and data preprocessing using Python, R, and Jupyter to enhance model performance and insights.
- Familiarity with **Hadoop components** including **HiveQL**, **Pig**, **HBase**, and **MapReduce** for querying, processing, and statistical methods to evaluate AI impact.
- Hands-on experience with **A/B testing** and experiment-driven decision-making, using conversion rate, lift analysis to optimize AI-driven business decisions.

- Strong proficiency in **containerization** and infrastructure as code Terraform, and **AWS CloudFormation** for creating scalable, automated cloud-native environments.

TECHNICAL SKILLS:

Category	Tools & Technologies
Generative AI & NLP	LLMs (BERT, T5, LaMDA, BLOOM), LangChain, GANs, Diffusion Models, spaCy, NLTK
Machine Learning & AI	XGBoost, Random Forest, SVM, PyTorch, TensorFlow, Scikit-learn, AutoML, Speech/Audio Processing, Voice Cloning, Speech Recognition
Explainable AI	SHAP, LIME
Cloud Platforms	GCP (Vertex AI), AWS (SageMaker), Azure (Synapse, AKS, Scale Sets)
Streaming & Messaging	Kafka, Apache Beam, Spark Streaming, Flume
Data Engineering & ETL	Apache NiFi, Airflow, Azure Data Factory (ADF), Azure Databricks, Oozie, Sqoop
Big Data Ecosystems	Hadoop, Spark, Hive, Pig, HBase, MapReduce
Databases & Warehousing	Snowflake, SQL Server, Azure Synapse Analytics, EDW
Programming Languages	Python, R, SQL, T-SQL, Java
Visualization & BI	Tableau, QlikView, Python (Matplotlib, Seaborn)
Version Control & DevOps	Git, CI/CD, Kubernetes, Docker
Other Tools	Jupyter Notebook, Pinecone (Vector DB), ELK Stack

PROFESSIONAL EXPERIENCE:

Client: Universal Music Group **Sep 2023 to Present**
Role: Gen AI Engineer

Roles & Responsibilities:

- Developed Generative AI and **TTS** models integrating LLMs with **GANs**, **Diffusion Models**, and **vocoders** (**Tacotron**, **WaveNet**) using Python and PyTorch for music creation, audio enhancement, and real-time style transfer.
- Implemented **transformer-based models** (**BERT**, **T5**) with **LangChain** and **LLM agents** to automate **metadata tagging**, **lyric summarization**, and **fan engagement chatbots**, improving efficiency in content delivery.
- Fine-tuned **LLMs** (**LaMDA**, **BLOOM**) to produce **artist biographies**, **marketing narratives**, and **personalized fan communications**, enabling rapid campaign deployment.
- Designed and deployed **semantic search pipelines** using **embeddings** and **Pinecone** for improved **music catalog discovery**, reducing retrieval times for internal and partner-facing systems.
- Implemented **speech recognition** and **voice cloning** workflows with **Librosa** and **ESPnet**, enabling multilingual transcription, high-fidelity artist voice replication, and rapid prototyping of conversational AI features.
- Optimized **real-time inference** for audio models, reducing latency to under one second and enhancing end-user interaction across music discovery and voice assistant platforms.

- Architected and deployed scalable **Gen AI applications** and **ML workflows** using **GCP Vertex AI** and **Kubernetes**, enabling production-grade, low-latency AI services for music and media workflows.
- Built **predictive analytics models** in **Python** and **PyTorch** to forecast streaming trends, optimize release schedules, and guide marketing strategy.
- Created **custom Python modules** for embedding generation, semantic search integration, and music metadata enrichment.
- Implemented **Python Flask APIs** to expose AI services, integrating with music analytics and distribution platforms.
- Designed and implemented **fraud detection systems** leveraging **statistical anomaly detection** and **machine learning models** to identify irregular streaming patterns, protecting revenue streams.
- Applied **Explainable AI techniques** (**SHAP**, **LIME**) to interpret **AI-driven recommendation algorithms**, ensuring stakeholder trust and transparency.
- Developed and exposed **RESTful APIs** using **Flask** to integrate AI capabilities into **music analytics** and **distribution platforms**.
- Conducted **A/B testing** on AI-generated **playlist recommendations** and **content personalization**, leading to measurable increases in listener engagement.

Key Tools and Skills : LLMs (BERT, T5, LaMDA, BLOOM), LangChain, GANs, Diffusion Models, Pinecone, GCP Vertex AI, Kubernetes, CI/CD, SHAP, LIME, Python, TensorFlow, PyTorch, Flask, REST APIs, Vertex AI Pipelines, A/B Testing, Jupyter, ELK Stack.

Client: SITA (Society International of Telecommunications and Aeronautics) **Oct 2021 to Aug 2023**
Role: AI/ML Engineer

Roles & Responsibilities:

- Built predictive models with **XGBoost**, **SVM**, and **Random Forest**, applying statistical methods and classification algorithms to enhance accuracy and model performance.
- Developed and deployed NLP solutions for text classification, sentiment analysis, and named entity recognition using Python (**NLTK, spaCy**) and transformer-based models, improving document processing efficiency.
- Explored **audio data preprocessing** and **dataset augmentation** using **Librosa** and **Kaldi** to support experimentation in speech recognition and conversational AI pipelines.
- Designed and implemented machine learning workflows in **GCP Vertex AI**, leveraging AutoML and custom models with feature engineering and hyperparameter tuning.
- Applied **deep learning techniques** using Neural Networks and PyTorch to perform complex predictive analytics and enhance model capability in processing unstructured data.
- Automated **end-to-end ML pipelines** using Vertex AI Pipelines, streamlining model training, deployment, and monitoring to support scalable and production-ready machine learning workflows.
- Built real-time analytics pipelines using **Kafka** and **Spark** Streaming to process streaming data and power operational dashboards for business-critical use cases.
- Integrated model explainability techniques such as **SHAP** and **LIME**, providing stakeholders with transparent insights into model predictions for regulatory compliance.
- Integrated Snowflake into ML workflows to streamline data ingestion and preparation, accelerating analytics and downstream model training.
- Developed and maintained Flask APIs and deployed ML models on **Kubernetes** within GCP, streamlining inference integration with enterprise applications.

Key Tools and Skills : XGBoost, SVM, Random Forest, NLTK, spaCy, Transformers, Vertex AI, AutoML, PyTorch, Python, Vertex AI Pipelines, Kafka, Spark Streaming, SHAP, LIME, Flask, Kubernetes, Snowflake, CI/CD.

Client: SSM Health **Feb 2018 to Aug 2021**
Role: Data Scientist/Engineer

Roles & Responsibilities:

- Designed and developed **ETL pipelines** using **Apache NiFi** and **Airflow** to process EHR data, radiology images, and clinical datasets, ensuring compliance with HIPAA standards.
- Optimized big data pipelines using **Hadoop**, **Spark**, and **Hive** for large-scale medical analytics, enabling population health studies and care quality monitoring.
- Built and maintained cloud-based ETL workflows in **Azure Data Factory (ADF)** for secure ingestion into **Azure Synapse Analytics**, **Azure Databricks**, and **Azure Data Lake**, improving data accessibility and streamlining end-to-end healthcare analytics.
- Developed predictive models in **Python** and **PySpark** for patient readmission risk, hospital resource allocation, and clinical outcome forecasting.
- Designed **Python ETL scripts** for healthcare data ingestion, cleaning, and feature extraction.
- Leveraged **Azure Databricks** and PySpark for advanced analytics on structured and unstructured healthcare data, including clinical notes and imaging metadata.
- Implemented data governance measures including **encryption**, **masking**, and **anonymization**, ensuring regulatory compliance.
- Created real-time streaming pipelines using **Kafka**, **Apache Beam**, and **Spark Streaming** for continuous patient monitoring and early alerting.
- Built interactive dashboards in **Tableau**, **QlikView**, and Python (Matplotlib, Seaborn) to support clinical decision-making and operational performance tracking.

Key Tools and Skills : Apache NiFi, Airflow, Hadoop, Spark, Hive, Azure Data Factory (ADF), Azure Synapse Analytics, Azure Databricks, PySpark, SQL, T-SQL, Data Lake, AKS, Azure Scale Sets, Kafka, Apache Beam, Tableau, QlikView, Python (Matplotlib, Seaborn), Healthcare Data Governance.

Client: Huntington Bank **Apr 2016 to Jan 2018**
Role: Data Analyst

Roles& Responsibilities:

- Performed **data cleansing, transformation, and exploratory data analysis (EDA)** using **Python, R, and Jupyter**, improving data readiness for reporting and analytics.
- Collaborated with **BI team** to gather reporting requirements and streamlined data ingestion by exporting data into **HDFS** and **Hive** using **Sqoop**, enabling faster analytics.
- Built **statistical modeling pipelines in Python** for customer segmentation and profitability analysis.
- Built **MapReduce** programs in Java to automate data preprocessing and transformation, significantly enhancing data quality and readiness for analysis.
- Wrote **Python automation scripts** for preparing datasets before Hadoop ingestion.
- Streamlined real-time data ingestion with **Flume** and optimized Hadoop ecosystem components, including **Pig**, **Hive** and **HBase**, reducing latency and enhancing platform stability and Performance.
- Developed **HiveQL** queries for business intelligence and trend analysis, enabling faster comparisons of current and historical data while improving **EDW** storage efficiency.

- Automated ETL pipelines with **Oozie** to simplify data ingestion and transformation, significantly improving workflow efficiency.

Key Tools and Skills : R, Python, Jupyter, Hive, HiveQL, HBase, Pig, Flume, Sqoop, Oozie, MapReduce (Java), EDW, HDFS, Statistical Methods, Text Mining, Cluster Analysis, Market Basket Analysis, Decision Trees.