Vijay Muni Reddy Mopuru
+1 973-449-0119 | vijayreddymopuru@gmail.com | LinkedIn | GitHub | NY, USA.

SUMMARY

- Results-driven Data Scientist with over 4 years of experience in designing, implementing, and managing scalable data solutions across diverse technological environments. Proficient in building and optimizing relational and non-relational databases, with expertise in performance tuning, data pipeline development, and ETL workflows. Skilled in cloud-based database services, including AWS RDS, Redshift, and Azure SQL Database, and adept at leveraging BigQuery and Snowflake for modern data warehousing solutions.
- Proven experience in building and deploying **machine learning (ML) and deep learning (DL) models**, leveraging frameworks like **TensorFlow, PyTorch, and Scikit-Learn** for predictive analytics, image processing, and natural language understanding. Skilled in data preprocessing, feature engineering, and model optimization to deliver impactful solutions in real-world business scenarios, such as fraud detection and risk management.
- Expertise in designing, implementing, and deploying **end-to-end ML pipelines**, including data ingestion, preprocessing, model training, and deployment using tools like **AWS SageMaker, Azure ML, and MLflow**. Successfully integrated ML solutions into production workflows, driving **automated decision-making** and **business intelligence**.
- Specialized in **Prompt Engineering** for **Large Language Models (LLMs)**, enabling advanced solutions for content generation, data summarization, and intelligent automation. Proficient in utilizing tools like **LangChain, LlamaIndex, OpenAI GPT**, and frameworks for **Retrieval-Augmented Generation (RAG)** to enhance model performance and optimize user-specific outcomes. Demonstrated ability to fine-tune LLMs for improved accuracy and business alignment
- Strong knowledge of SQL, PL/SQL, and Python, with experience in writing complex queries, stored procedures, and automating workflows using tools like Ansible, PowerShell, and Apache NiFi. Demonstrated ability to design and maintain data platforms that ensure high availability, reliability, and scalability.
- Designed and developed scalable **ETL pipelines** to process and transform large datasets, ensuring high data integrity and performance optimization.
- Designed and conducted large-scale statistical experiments to optimize hiring pipelines, improving candidate experience and reducing time-to-hire by 15%.
- Developed predictive models leveraging machine learning algorithms (e.g., random forests, gradient boosting) tailored to enhance hiring efficiency and decision-making processes.
- Collaborated with cross-functional teams, including product managers, engineers, and scientists, to deliver data-driven solutions that address business challenges and improve operational efficiency.
- Implemented data preprocessing and feature engineering techniques using Python, Pandas, and NumPy to prepare large datasets for machine learning workflows.
- Led end-to-end ML model development projects, including requirement gathering, algorithm selection, hyperparameter tuning, and production deployment, to optimize hiring analytics.
- Engineered robust **data models and warehouses**, enabling efficient storage and retrieval of structured and unstructured data for analytics and reporting purposes.
- Implemented **AWS-based solutions**, utilizing services like **Redshift**, **S3**, **Glue**, **EMR**, **Kinesis**, and **Lambda** to deliver scalable, cloud-native data pipelines.
- Skilled in crafting and optimizing **AI prompts** for **Generative AI models**, enabling improved accuracy and contextual relevance in solutions like **content creation, automated reporting, and intelligent search engines**. Integrated **LangChain** and **LlamaIndex** for dynamic prompt chaining and retrieval-augmented generation (RAG) workflows.
- Developed and automated complex **dashboards and reporting systems** to provide actionable insights for driving key business decisions.
- Hands-on experience in Machine Learning and Deep Learning concepts, leveraging frameworks such as TensorFlow and PyTorch for implementing predictive models and pipelines. Skilled in data preprocessing, feature engineering, and model optimization, enabling seamless integration of ML capabilities into data workflows.

PROFESSIONAL EXPERIENCE

Data Scientist Dot IT Sol **Jan 2024 – Till Date**

- Design and Maintain Scalable Data Pipelines: Architect, implement, and optimize ETL/ELT workflows for ingesting, transforming, and delivering financial datasets, leveraging cloud platforms like AWS (Redshift, S3) or Azure, ensuring data accuracy, reliability, and scalability.
- Support Advanced Analytics and ML Initiatives: Collaborate with data scientists to preprocess large-scale banking datasets, perform feature engineering, and deploy ML/DL models using frameworks like TensorFlow and PyTorch to enhance fraud detection and risk management systems.
- Developed and deployed advanced **Deep Learning models** for **computer vision** (e.g., CNN-based image classification) and **natural language processing** tasks, utilizing frameworks like **Keras** and **Hugging Face Transformers**. Improved accuracy in real-world use cases, such as **brain tumor detection** and **driver activity monitoring**.
- Collaborated with stakeholders to identify business requirements, translating them into ML solutions to enhance decision-making processes, such as fraud detection and risk management systems at Bank of America.
- Built predictive models to improve operational efficiency, using TensorFlow and PyTorch for regression and classification tasks.
- Develop and Optimize Data Models: Design and implement relational and dimensional data models for data warehousing and analytics, utilizing tools like Snowflake, Big Query, or Hive, to support regulatory compliance and real-time reporting.

- Designed and deployed AI-driven automation solutions, enhancing **business process efficiency** through intelligent systems for fraud detection, anomaly detection, and demand forecasting. Combined **AI and statistical techniques** to deliver actionable insights that improved decision-making.
- Automate Data Workflows: Implement and manage automation tools like Apache Airflow and Python-based scripts to streamline data processing, integrate ML/DL models into production pipelines, and ensure compliance with internal governance and security standards.
- Enhance Data Insights and Decision-Making: Build real-time dashboards and predictive analytics tools using Tableau, Power BI, or custom Python visualizations, empowering stakeholders to make data-driven decisions for customer experience improvements and operational efficiency.
- Designed and deployed automated data pipelines using ETL frameworks to streamline data collection, integration, and processing, resulting in a 20% reduction in manual effort.
- Conducted exploratory data analysis (EDA) to identify trends and patterns in hiring data, providing actionable insights that enhanced the associate experience.
- Researched and implemented state-of-the-art machine learning algorithms to forecast hiring trends and optimize workforce planning strategies.
- Fostered a culture of continuous improvement by mentoring junior data scientists and engineers, ensuring adherence to best practices in statistical modeling and data engineering.
- Integrated advanced natural language processing (NLP) techniques to analyze candidate feedback and improve recruitment communications.
- Created and maintained detailed documentation for machine learning models, statistical experiments, and analytical workflows to ensure transparency and reproducibility.

Data Engineer Intern – Dot IT Sol., United States **May 2023 –Sep 2023**

- Reduced processing time by 80% and increased data accuracy of deep learning models by 10% by developing a data preprocessing pipeline for Data gathering and cleaning precipitation data from NASA.
- Achieved an accuracy rate of 65% by researching and implementing cutting-edge machine learning algorithms to develop SRGAN models for transforming low-resolution images into high-resolution images.
- Proficient in fine-tuning **Large Language Models (LLMs)** such as GPT, LLaMA, and **BERT** for domain-specific tasks, including chatbots, document summarization, and code generation. Optimized models using techniques like **PEFT** (Parameter-Efficient Fine-Tuning) and **Knowledge Distillation** to enhance performance with minimal resource usage.
- Performed Statistical analysis in python and Exploratory Data Analysis on precipitation data to extract key features and provide Analytical solutions, resulting in a 30% improvement in image resolution accuracy.
- Collected, pre-processed, and analysed large datasets for machine learning models, ensuring data conformity to ML algorithm assumptions**.**
- Developed custom statistical methodologies to evaluate the performance of hiring strategies, resulting in improved ROI on recruitment initiatives.
- Presented findings and recommendations to senior leadership, translating complex technical concepts into actionable business insights to guide strategic decision-making.
- Implemented techniques for **model monitoring, retraining**, and performance optimization in production environments, ensuring models remain accurate and aligned with evolving business requirements. Utilized tools like **TensorBoard** and **Weights & Biases** for real-time metric tracking and model validation.
- Developed and deployed ML/DL models for cryptocurrency price forecasting, leveraging LSTM neural networks to reduce error rates by 15%.
- Designed SRGAN models for enhancing image resolution, improving performance accuracy by 10% during data analysis projects.

Graduate Teaching and Research Assistant - NJIT **Jan2023–May2023**

- Participated in faculty meetings to discuss course objectives and updates, ensuring alignment with academic goals.
- Designed, evaluated, and graded assignments and projects, providing detailed feedback to enhance student understanding and performance.
- Conducted lab sessions and office hours, offering hand-on guidance and support to students on AI.
- Collaborated with faculty on research projects, assisting in data collection and analysis to support innovative AI research initiatives.
- Facilitated group discussions and study sessions, fostering a collaborative learning environment that encouraged student engagement and critical thinking.

Data Engineer - Spring Info Tech Private Limited **Nov 2020 - Jul2022**

- Developed predictive models for cryptocurrency price forecasting using linear regression, decision trees, and random forests.
- Enhanced prediction accuracy by integrating LSTM neural networks for Ume-series forecasting, reducing model error rates
- Utilized Azure services for data integration workflows, contributing to a 35% optimization of ETL processes, and created PowerBI dashboards to track KPI metrics, reducing manual data analysis by over 50%.
- Generated Azure monitoring dashboards minimized customer pain points by 20%, & supported analysts with troubleshooting, ensuring seamlessness
- Streamlined data ingestion and processing using Azure Databricks, and developed Spark applications for data extraction and transformation, enhancing overall data warehousing efficiency.

- Utilized PySpark on Databricks to enhance ETL process, reducing processing time by 30% for large datasets, effectively improving query efficiency and database performance.
- Leveraged automation tools such as Ansible, Python, and PowerShell to streamline database deployment, configuration, and maintenance, reducing manual effort and error rates.
- Created and implemented scripts for automated database backups, monitoring, and performance tuning, improving overall operational efficiency.

Data Engineer Intern - Spring Info Tech Private Limited **Jan 2021 – Sep2021**

- Created a handwritten digit recognition system using CNN and SVM algorithms with the MNIST dataset, achieving a 13% improvement in accuracy.
- Developed and maintained robust database solutions using PostgreSQL, MySQL, Amazon Redshift, and AWS RDS, optimizing data storage and retrieval processes.
- Optimized feature extraction and hyperparameter tuning to enhance model performance, evaluating results with metrics like accuracy, precision, recall, and F1 score.

Web Development Intern - The Spark Foundation **Sep 2020 – Dec 2020**

- Create Visually Appealing Designs: Develop engaging and responsive designs for websites and web applications using design tools like Adobe XD, Figma, and Sketch.
- User Experience (UX) and User Interface (UI) Design: Conduct user research and create, wireframes, mock-ups, and prototypes for enhancing.

SKILLS

Category	Skills
Languages	Python (Pandas, NumPy, PySpark) Flask, R, C/C++, SQL, Java
Frameworks	TensorFlow, Keras, PyTorch, CUDA, OpenMP, REST, Langchain, Celery, Pandas, NumPy, Scipy, Sklearn.
Data Engineering and Analytics	MongoDB, MySQL, PostgreSQL, Pinecone, Azure Storage Service, RDS, ETL, Athena, Kinesis, S3, Spark, Data Analytics
ML/AI	Machine Learning, Deep Learning, ML Pipeline Orchestration, Generative Adversarial Networks (GANs), Computer Vision, Natural Language Processing (NLP), Large Language Models (LLMs), LLM Fine-tuning, Quantization, Knowledge Distillation, Retrieval Augmented Generation (RAG), Transformers, PEFT, OpenAI, Llama, Gemini Pro, Hugging Face, Bedrock, Sage Maker, Distributed LLM Training, RLHF, Data Warehousing, Hypothesis
Tools and Platforms	Git & GitHub, Kubernetes, Docker, OpenShift, Azure DevOps, AWS, Azure,
Cloud Platforms	AWS(S3, EC2, Lambda, EC2, DynamoDB) Azure (Azure SQL, Cosmos DB, Azure Data Lake, Synapse Analytics), Google Cloud (BigQuery, Data Catalog), Salesforce (System Analysis)
BI & Analytics	Power BI, Tableau Spark SQL, MLflow (Machine Learning Operations)
Software Development	Backend Development, API Optimization, Parallelizing with Concurrent Futures Understanding of Common RDBMS, Experience with Clustering and SAN Infrastructure

Certifications:

AWS Certified Data Engineer Ass, Microsoft Certified Azure Data Engineer Associate, Microsoft Certified Power BI data Analyst Associate, Azure Fundamental, Generative AI Accreditation from Databrick, Python from Microsoft, SQL from Microsoft, Joy of Computing with Python by NPTEL, Azure Databricks & Spark foundation, Data Engineering on Microsoft Azure, Intellectual Property Rights.

PUBLICATIONS:

Cryptocurrency Price Prediction Research:

- Developed a real-time and adaptive cryptocurrency price prediction model utilizing Twitter sentiments and Reddit community analyses to forecast price movements for Bitcoin and Ethereum, leveraging both time series analysis and social media data.
- Conducted an exploratory analysis of social media sentiment correlations with cryptocurrency prices, enhancing prediction accuracy and offering insights into price fluctuations for Bitcoin, Ethereum, and Binance.

- Demonstrated expertise in researching cutting-edge **Generative AI** trends and applying solutions like **multi-modal models** and advanced **NLP techniques** to solve complex business challenges. Published insights on **cryptocurrency price prediction** and social media sentiment analysis using time series and ML models.

NOTABLE PROJECTS

LinkedIn - Automated Job Application Tool **Jun 2024 – Aug 2024**

- Created an AI-driven tool to automate job searches and applications on LinkedIn, featuring customizable search criteria, one-click application submissions, and dynamic resume generation.
- Developed a Python-based job application automation tool that leverages YAML for configuration management, streamlining the application process for LinkedIn jobs.
- Designed and implemented data models using Python's `data class`, enhancing code readability and maintainability while capturing essential candidate information, including self-identification, legal authorization, work preferences, availability, and salary expectations.
- Integrated intelligent job search automation and AI-powered personalization to streamline the application process, secure data handling, and ensure efficient profile management, Configured and installed the tool using python and YAML files for dynamic application.

Brain Tumour Detection from MRI Data Using CNNs **Aug 2023 – Dec 2023**

- Developed a Convolutional Neural Network (CNN) model to detect brain tumors from MRI images, utilizing a dataset of 3,762 MRI scans and 13 extracted features.
- Addressed challenges such as class imbalance and tumour variability by implementing binary classification and evaluating model performance using metrics like AUC-ROC, Sensitivity, and F1-Score.
- Designed and implemented a Convolutional Neural Network (CNN) architecture to detect and classify brain tumors with an accuracy of 93%.
- Developed a pipeline for segmentation and localization of brain tumors using advanced deep learning techniques.
- Applied transfer learning using pre-trained models (e.g., VGG16, ResNet) to improve classification performance and reduce training time.
- Leveraged Grad-CAM and other explainability tools to analyze CNN decision-making and improve model interpretability for medical experts.
- Built an end-to-end deployment pipeline for the detection system using Flask and Docker, allowing seamless user interaction via a web interface.
- Utilized TensorBoard for real-time monitoring of training metrics such as loss, accuracy, and validation performance.

Driver Activity Detector **Jan 2022 – May2022**

- Designed a driver monitoring system to detect drowsiness through eye blinking and head movements using camera data, achieving over 80% accuracy.
- Employed logistic regression models with one-vs-one binarization to classify drowsiness states, enhancing traffic safety.
- Processed and preprocessed video and sensor data by implementing frame extraction, normalization, and augmentation techniques, enhancing model accuracy.
- Trained a Convolutional Neural Network (CNN) to classify driver activities such as hands on wheel, texting, eating, or drowsiness with an accuracy of 90%.
- Integrated data from multiple sensors (camera, accelerometer, and gyroscope) to create a multi-modal detection system for robust activity classification.
- Applied transfer learning with pre-trained models (e.g., MobileNet, ResNet, YOLO) to reduce training time and improve detection performance Utilized OpenCV for motion tracking and pose estimation to identify driver postures and detect distractions.
- Developed and deployed the system on edge devices (NVIDIA Jetson Nano, Raspberry Pi) for real-time in-vehicle processing.
- Implemented real-time data streaming pipelines using Kafka and MQTT, ensuring seamless data collection and processing.
- Devised a drowsiness detection mechanism using facial landmarks and eye aspect ratio analysis to alert drivers in potentially dangerous situations.

Education

New Jersey Institute of Technology, Newark, NJ August 2022 – December 2023
Master of Science in Machine Learning, GPA – 3.6
Nitte Meenakshi Institute Of Technology, Bangalore, India July 2018 – September 2022
Bachelor of Engineering in Computer Science, GPA – 8.0