Raghavendra Chunduri PhD

Data Science MLOPs With Generative AI

rajasekhar.a@meridiansoftt.com LinkedIn

**Phone: 380-388-3682, Work Authorization: GreenCard**

SUMMARY

Highly accomplished Data Scientist and MLOPs Engineer with a Ph.D. in Scalable Machine Learning and over 14 years of combined expertise in Data Science, MLOPs, Generative AI, and Machine Learning.

Specialized in developing advanced AI-driven solutions leveraging large language models (LLMs), deep learning, and generative AI techniques. Proven ability to build and deploy end-to-end ML pipelines, NLP models, and Generative AI applications to address complex business challenges and drive decision-making.

Expert in implementing scalable, cloud-based AI solutions using platforms like Azure and AWS. A strong collaborative leader who brings innovative, high-impact data science and MLOPs strategies to life in real- world applications across various industries

# Skills

* **Cloud Platforms:** Microsoft Azure (Azure Machine Learning, Cognitive Services), AWS, GCP
* **Machine Learning & AI:** Supervised/unsupervised learning, Large Language Models, NLP, Deep Learning
* **Tools & Libraries:** TensorFlow, PyTorch, scikit-learn, pandas, NumPy, Spark, Big Query, Databricks, Keras
* **Programming Languages:** Python, R, SQL, Scala, Java, C++
* **AI Model Deployment:** Azure ML Studio, Vertex AI, Sage maker
* **Data Processing:** Hadoop, Apache Spark, HDFS, Hive, Pig, Azure Synapse Analytics
* **Big Data Tools:** HDFS, MapReduce, Spark, Kafka, Flume, Oozie
* **Generative AI:** Experience in applying Generative AI (GANs, VAEs) for text generation, automated content creation, and virtual agent development.
* **Brain-Aided Computing:** Advanced knowledge in integrating brain-computer interface (BCI) technology with AI for cognitive enhancement and human-AI collaboration.
* **Data Science & Analytics:** Skilled in data wrangling, feature engineering, statistical analysis, and model validation for AI projects.
* **Automation & NLP:** Proficient in building automated text analysis, chatbots, and sentiment analysis tools using natural language processing.

#  Certifications

* Microsoft Certified: Azure Data Scientist Associate
* Microsoft Certified: Azure AI Engineer Associate
* Certified Hadoop, Apache Spark Developer
* Stanford University: Machine Learning Certificate
* Coursera: Hadoop Platform and Application Framework – University of California, San Diego

# Educational Qualifications

## Postdoctoral Research Fellow –

Neuromorphic Computing

Department of Electrical and Computer Engineering, University of Colorado, Colorado Springs

* + **PhD** – Scalable Machine Learning

School of Computer Science Engineering, VIT University, India

* + **Master of Technology** – Computer Science and Engineering

JNTU Kakinada, India

* + **Bachelor of Technology** – Computer Science and Engineering

JNTU Kakinada, India

# Research Projects

## Project: Like2Vec

Technologies: **Azure Databricks**, **Apache Spark**, **Python**

Developed a recommender system based on **word2vec** for transforming sparse datasets into dense latent spaces for high-accuracy recommendations in a production environment using **Azure** services.

## Project: Isomantics NLP

Technologies: **Azure ML**, **Python** Explored mapping neural word embeddings between languages using **NLP** and **Azure Cognitive Services** for cross-linguistic applications

# Professional Experience

## Union Pacific Railroad

**System Consultant – Data Science & AI/ML Engineering Omaha, NE** | **August 2019 – Present**

* Developed LLM-based chat summarization applications using **Vertex AI** and **Azure OpenAI** for OSS support, resulting in significant automation of internal processes.
* Designed and deployed a **GenAI-based data analysis engine** to bridge data discrepancies between legacy mainframe systems and modern transportation systems.
* Spearheaded the development of **ML applications** for predicting customer retention and basket value using

**PySpark** and **TensorFlow** on **Azure Databricks**, improving accuracy and model performance.

* Leveraged **Azure ML Studio** for hyperparameter tuning and optimization, achieving an 87% F1 score for retention models and 93% R-squared for basket value prediction models.
* Orchestrated large-scale **data engineering projects** on **Azure Databricks** using **PySpark**, **Hive**, and

**BigQuery**, improving customer experience scores by 5%.

* Led the development and deployment of **CX elasticity models**, driving an 18% increase in customer purchasing frequency.
* Technologies Used: **Azure Databricks**, **PySpark**, **LLMs**, **GenAI**, **MLOPs**, **BigQuery**, **Vertex AI**, **Azure ML Studio**, **Airflow**
* Implemented Optimization techniques like Regularization, Cross-validation, Hyperparameter Fine Tuning

improving the retention model’s F1 score to 87% and basket value model’s R-squared to 93%.

* Facilitated cross-departmental collaboration, resulting in 10% improvement in data accuracy and 20% increase in project efficiency.
* Transformed raw data into actionable insights, contributing to a 5% improvement in CX strategy effectiveness.
* Facilitated knowledge transfer and skill development, mentoring 3 junior team members to proficiency in data engineering tools and techniques.

## Kaiser Permanente

**System Consultant – Data Science Omaha, NE** | **June 2016 – July 2019**

* Designed and implemented real-time data pipelines using **Azure Databricks** for ingesting, cleaning, and transforming healthcare data from multiple sources.
* Deployed ML models in production environments for **clinical data validation** on **Azure ML**, leading to more accurate patient diagnosis.
* Built scalable data processing frameworks using **Apache Kafka** and **Azure Storage** solutions like **Blob Storage** and **Data Lake**.
* Technologies Used: **Azure Databricks**, **Azure ML Studio**, **PySpark**, **Kafka**, **Pandas**, **NumPy**, **Scikit- Learn**
* Collaborated with auditors and data engineers to identify AI use cases and opportunities, driving innovation and efficiency.
* Led development of a GenAI-powered chatbot for the financial sector using Python, NLP, and OpenAI's API, boosting user engagement and operational efficiency.
* Optimized vector databases (Chroma DB, Pinecone), enhancing data retrieval capabilities for AI applications by 30%
* Implemented text summarization using AWS tools, effectively condensing lengthy documents into concise summaries. Developed customized solutions for varied business needs, integrating advanced analytics to drive strategic decision-making and operational improvement

## T-Mobile - User Experience Project Software Test Engineer/Data Analyst Omaha, NE | November 2013 – May 2016

* Developing Selenium Web driver-JAVA-Maven-TestNG-POM-Data Driven-Keyword driven framework.
* Developed reporting dashboard to track test coverage.
* Working on Selenium WebDriver 3.0 using JAVA
* Used Jenkins for Continuous integration and GIT lab for code hoisting platform. Used various GIT.
* commands to edit, upload, commit and push the code to Gitlab.
* Developing test scripts, executing the developed Test Scripts, Debugging and Analyzing the Test Results.
* Executing the Test Scripts on AUT for test data generation.
* Performing API testing using Ready API, validating JSON responses, and automating the same using Ready API 2.4.0.
* Built real-time data processing systems using Apache Kafka and stream processing frameworks.
* Designed and implemented data storage solutions using cloud-based technologies such as AWS S3 .
* Developed data pipelines for ingesting, cleaning, and transforming data from multiple sources.
* Created data models and performed data analysis to support business intelligence and reporting needs.
* Developing generic web and business functions that can be reused in the framework.
* Used SQL queries to fetch test data required from excel sheet. Performed database validations.
* Working with Oracle SQL Developer to fetch data from database and perform data validations against application data. Performing Database migration testing for 14million customers.

## Union Pacific Railroad System Consultant – SDET

**Hyderabad India** | **November 2010 – Oct 2013**

* Successfully Implemented scalable data pipelines using Apache Spark and Java, Python to process large volumes of data, providing ongoing support and enhancements.
* Analyzed data and developed data queries for large data analytics and used Tableau reports for data validation, Grafana dashboards for performance monitoring.
* Data preprocessing utilizing Weka ETL tools and R
* Designed and optimized database schemas to ensure data integrity and performance.
* Collaborated with cross-functional teams to gather data requirements to provide insights for business decision- making.
* Implemented data quality checks and monitoring systems to ensure data accuracy and reliability.
* Participate in daily scrums, Backlog refinement, grooming, Sprint planning meetings, Sprint demo, and retrospectives.
* Managing the offshore team and providing strategy to improve data quality.
* Conceptualize and develop relevant use cases for proof-of-concepts (POC)
* Establish project budget based on the estimates to full fill project requirements.
* Participate in daily scrums, Backlog refinement, grooming, Sprint planning meetings, Sprint demo, and retrospectives.
* Managing the offshore team and providing strategy to cover both manual and automation testing.
* Handling multiple projects with team size of 10 manual and automation resources.
* Define overall program test strategy, automation strategy, entry and exit criteria.
* Review the Test Basis, design and document Test Plan, Test Scenario, Test Cases.
* Define test data strategy by analyzing business requirements, product types, Market segment impacted, data elements & enumeration in web services, APIs, daily Jobs.
* Active involvement in the design and implementation of automation framework for smoke and regression test suites.
* Conducted end-to-end testing of the SAP-based way billing system to ensure accuracy, reliability, and compliance with business requirements and meet revenue team standards.
* Developed comprehensive test plans and test cases to cover various billing scenarios, including invoicing, payment processing, and reporting.
* Automation test cases development using TOSCA for SAP based UI’s
* Be responsible for the writing and maintenance of flawless automation using IBM rational integration tester.
* Designing integration tests and implementing integration automation framework that covers different technologies like mainframes, UI, and middle ware applications.
* Designed the automation report dashboard to track failed and passed testcases/
* Create and maintain repository of known test defects.