

Ritesh Zambre

Pune, India | riteshzambre111@gmail.com | +91 7028224490 | linkedin.com/in/riteshz

Summary

Strategic and results-oriented Senior Data Scientist with an MBA in AI & Machine Learning and 5+ years of experience (including 1 year as intern) architecting and deploying impactful AI/ML solutions. Expertise in LLMOps, MLOps, Generative AI, and DataOps, driving business value across diverse industries, including healthcare, finance, and retail. Proven ability to lead end-to-end project lifecycles, from secure data engineering on cloud platforms (Azure, AWS, GCP) to deploying scalable models and BI solutions (Microsoft Fabric, Power BI, RShiny). Using advanced Python, R, PySpark and AI techniques to solve complex challenges and deliver measurable results is a skill.

Experience

- Software Engineer (Senior Data Scientist), e-Zest Solutions – Pune, India** Sept 2023 – Present
- Spearheaded secure data onboarding and transformation for **WHO** projects (Covid, TB, AMR, etc.) using Azure (DevOps, AML, ADF, Key Vault, Service Principals), processing the data, reduced onboarding time by almost 50%.
 - Architected and delivered a **comprehensive data analytics tool** for the GLASS AMR team, involving complex ETL, advanced RShiny visualizations, and xMart data storage, used researchers / analysts across the globe.
 - Designed & implemented CI/CD pipelines using GitHub Actions, ensuring streamlined and automated model deployment and versioning processes. Optimized **WHDH** data access libraries, AML workflows, achieving 50% faster data retrieval for critical health data.
 - Managed deployment and maintenance of 700+ WHO ShinyApps on Azure, ensuring high availability and performance. Developed a secure, reusable ShinyApp template with SSO (Azure Entra ID) and xMart integration, adopted by 30+ WHO development teams, reducing development time by 40%.
 - Leveraged PySpark in Azure Databricks for distributed machine learning model training and large-scale data processing, handling data sets up to 30 GB efficiently.
 - Orchestrated MLOps workflows on Azure for efficient deployment, monitoring, and management of machine learning models in production environments.
 - Spearheaded the development and deployment of a Generative AI-powered chatbot using Azure AI Studio, enabling real-time responses to user queries on pull requests (open, merged, latest PRs).
- Senior Data Scientist, Voiceback Analytics – Pune, India** Sept 2021 – Sept 2023
- Led a flagship data science project for **Theobroma (Food Retail)**, overseeing data engineering (Azure Data Factory ETL pipelines for Inventory, Sales, Stockout), Power BI dashboard development, and advanced analytics. Delivered actionable insights that [Quantifiable Impact: reduced stockouts by 7%, improved inventory turnover by 2%].
 - Implemented Market Basket Analysis, RFM segmentation, sentiment analysis (Python), and demand forecasting models, leading to [Quantifiable Impact: 20% improvement in forecast accuracy].
 - Developed and deployed 3 critical machine learning models (Python, Azure ML Studio) for a **Microfinance company** to predict loan defaults. Engineered features from diverse datasets, contributing to a [Quantifiable Impact: 0.36% reduction in credit losses] and forecasting 12-month expected credit losses (ECL).
 - Designed and implemented a rule-based hourly production planning model (Python) for **Paradise Food Court**, integrated with Power BI and Azure Functions. Enhanced with Power Apps for user input (shelf life, conditions), directly [Quantifiable Impact: e.g., reducing food wastage by 9%, optimizing production schedules].
- Data Science Intern (Unpaid), Innodatatics.ai – Pune, India** Dec 2020 – Sept 2021
- Developed and deployed a multinomial logistic regression model (Python, Heroku, Streamlit) to classify baggage operator performance for a leading beverage company.
 - Created MIS dashboards in Tableau & built data pipelines using AWS Glue, improving reporting efficiency.

Education

Dr. D. Y. Patil Vidyapeeth, Pune, MBA in Artificial Intelligence & Machine Learning
Savitribai Phule Pune University, Bachelor of Engineering

June 2022 – Sept 2024
Aug 2017 – Dec 2020

Skills

Programming & Databases: Python , R, PySpark, SQL, DAX, NoSQL, Vector Databases

Data Engineering & Big Data: Azure Data Factory (ADF), Azure Synapse Analytics, Azure Data Lake, Azure Databricks, AWS Glue, Apache Spark, Hadoop Ecosystem (HDFS) , Kafka, Data Warehousing (Kimball, Inmon), ETL / ETL Design, Data Modeling, Data Governance, API Integration (REST, SOAP).

Cloud Platforms: Microsoft Azure (Azure ML Studio, Cognitive Services, Azure Functions, Key Vault, Blob Storage, AKS), Amazon Web Services (AWS SageMaker, S3, EC2, Lambda, Redshift), Google Cloud Platform (GCP AI Platform, BigQuery) .

Machine Learning & Deep Learning:

- **Core ML:** Regression, Classification, Clustering, Dimensionality Reduction, Ensemble Methods (Random Forests, Gradient Boosting), SVM, KNN, Naïve Bayes, Time Series Analysis (ARIMA, Prophet).
- **Deep Learning:** Neural Networks (ANN, CNN, RNN, LSTMs), Transformers (BERT, GPT models), Computer Vision (Image Classification, Object Detection), NLP (Text Classification, Sentiment Analysis, NER, Topic Modeling).
- **Advanced Topics:** Reinforcement Learning, Anomaly Detection, Recommendation Systems, Statistical Modeling, Experimental Design.

Generative AI & LLMOps:

- **LLMs:** OpenAI GPT-series, Llama, BERT. Prompt Engineering, Fine-tuning, Retrieval Augmented Generation (RAG).
- **Frameworks & Tools:** LangChain, LlamaIndex, Hugging Face Transformers, Vector Databases (e.g., Pinecone, Milvus, ChromaDB).
- **LLMops:** LLM deployment, monitoring, versioning, ethical AI considerations for LLMs.

MLOps & DataOps: CI/CD for ML (Azure DevOps, Jenkins, GitHub Actions), Model Deployment & Serving (Docker, Kubernetes, Azure Container Instances, SageMaker Endpoints), Model Monitoring & Management (MLflow, Kubeflow) , Feature Stores, Data Versioning (DVC), Workflow Orchestration (Airflow, Azure Data Factory).

BI, Visualization & Reporting: Power BI, Tableau, RShiny (Shinyapps.io).

DevOps & Other Tools: Git, Docker, Kubernetes, Heroku, Streamlit, Linux/Unix, JIRA, Confluence.

Professional Attributes: Strategic Thinking, Advanced Analytical & Problem-Solving Skills, Project Leadership & Management, Stakeholder Communication & Management, Cross-functional Team Collaboration, Business Acumen, Adaptability, Mentoring.

Domain Expertise: Healthcare Analytics, Financial Services (Credit Risk, Fraud Detection), Retail & CPG (Demand Forecasting, Inventory Optimization, Customer Segmentation), Supply Chain.

Certifications

- | | |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none"> • Udemy – MLOps, GitHub Copilot • Udemy – Generative AI, Hugging face, LangChain • HackerRank, IBM - Python, SQL | <ul style="list-style-type: none"> • 360DigitMG – ML & DL, Tableau Visualizations • Gen AI Advanced • ineuron – Power BI, Microsoft Fabric |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Achievements and Awards

- Successfully reduced out-of-stock items by 7% and returns of short shelf-life products by 4% for a major food retail client, directly impacting revenue by 16 Cr. INR/ 2 % and significantly aiding area managers in demand planning accuracy.
- Played a vital role in the migration of WHO Power BI workspaces to Microsoft Fabric in Collab with Microsoft

Self Paced Learning

- | | |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none"> • Enterprise RAG-Based GenAI Knowledge Assistant • Multi-Agent AI Task Automation System • End-to-End MLOps Pipeline with CI/CD Monitoring • LLM Fine-Tuning Evaluation Framework | <ul style="list-style-type: none"> • Real-Time ML Inference with Drift Detection • AI-Powered Text-to-Insights Data Analyst • Kubernetes-Native Model Serving Platform • Autonomous GenAI Research Report Generator |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|