

# ARYA VEER KRISHNA

## Data Scientist

### WORK EXPERIENCE

#### JIO Platforms Limited — Mumbai

Data Scientist 12/2023 – Present

- Built a **Scalable Generative AI Platform** with Observability, Deployment, Data, and Compute + Foundation Layers, enabling **onboarding of Reliance and Jio businesses into Gen-AI within 7 days** with minimal stack knowledge.
- Delivered **Modular RAG chatbots** as part of the Jio Analytics COE team, reducing LLM calls by **25%**, cutting response time by **40 seconds**, and saving approximately **₹1.25 Lac/month** through query routing and LLM-based pipeline evaluation.
- Engineered **data ingestion pipelines** using Ray to process and store **~9.8 Mn** unstructured proprietary data in vector databases.
- Developed **graph extraction pipelines** for PDF data, achieving **92% accuracy** in converting bar/line graphs to CSV files using LLMs.
- Built **summary indexing pipelines** for Wiki, HR, Git, RAG-to-SQL/HIVE, and DevOps use cases, increasing retrieval accuracy by **23%** in chatbots and multi-agent systems using Milvus.
- Designed **CDC pipelines** to track Git and Wiki changes, ingesting **~10K repository data** into vector databases.
- Created **multi-agents** using Langgraph to perform root cause analysis (RCA) for Logi Composer errors, resolving issues in **~6 minutes** and safeguarding **~9.7 Mn** live data records daily.

### PROJECTS

#### Customer Churn Prediction using Chaid & LLM for Insights 11/2024

- Engineered a predictive analytics framework leveraging CHAID and Gemini LLM, which streamlined churn analysis processes and increased reporting efficiency by 40%, allowing for quicker response to customer trends.

#### Chart-to-CSV 04/2024

- This project extracts images from PDF files and runs inference using the DePlot model to generate data tables from graph images with 92% accuracy on Simple Bar/line

charts.

- The setup uses Docker and Docker Compose to simplify deployment and make it easy to start.

### CONTACT

• Mumbai, India (Open to Remote) • +91-9122112912  
• krishnaaryaveer@gmail.com • [linkedin](#) • [portfolio](#)

### SKILLS

#### Main Skills & Tools:

- Python • PySpark • Pandas • Statistics • LangChain • Milvus
- Ray Framework • Computer Vision • Kubernetes • Docker
- Scalable Solutions • DSPy • Prompt Engineering • LLMs • Modeling
- SQL
- Kafka
- Predictive Analytics • Machine Learning (ML)
- CHAID
- Research
- Natural Language Processing (NLP)

#### Complementary Skills and Tools:

- Airbyte • Bash • Redis
- Airflow
- Kafka
- Pandas
- Git
- PySpark

### EDUCATION

#### Bachelors of Technology in Electronics and Telecommunications Engineering

K . J . Somaiya Institute of Technology  
Mumbai — 10/2019

GPA: 9.32/10

### OTHER

#### Certifications:

- Google Python Programming Course: Coursera.
- Google Cloud: Machine Learning Operations (MLOps): Coursera

#### Accomplishments:

- Published an IEEE paper in the 2022 5th International Conference on Advances in Science and Technology (ICAST) titled "A Modern Approach to Conventional Silk Farming" (DOI: 10.1109/ICAST.2022.100395).