

Dhiraj Johare

dhirajjohare@outlook.com | +91-8879883952 | Mumbai, India
linkedin.com/dhirajjohare | kaggle.com/therealdj97 | github.com/therealdj97 | huggingface.co/therealdj97

AI and Big Data Engineer with 2+ years of experience developing NLP systems, LLMs, and predictive analytics. Skilled in enhancing chatbot performance, deploying RAG pipelines, and optimizing backend architecture for scalability. Proven ability to automate CI/CD processes and improve forecasting accuracy. Strong collaborator with a focus on delivering AI-driven solutions and operational efficiency.

Experience

AI Engineer, Tech Mahindra – Mumbai, MH, India Aug 2024 – Present

- Accomplished successful POCs of multiple AI use cases to improve user experience.

AI & Big Data Engineer, ShikshaML (DreamsAI Innovations) – Mumbai, MH, India Sept 2023 – Jun 2024
Founder & Managing Director's office

- Accomplished the design and development of Large Language Model (LLM) and Retrieval Augmented Generation (RAG) based NLP systems with advanced prompt engineering techniques, ensuring alignment with industry trends and standards while utilizing Langchain, Faiss, GPT-3.5 turbo, GPT-4 turbo, Whisper, and Pinecone.
- Set up an on-premise multi-GPU NVIDIA node for the Llama3 LLM model with a load balancer, process manager, and reverse proxy, improving chatbot response time by 30% and increasing RAG efficiency by 25%.
- Collaborated with cross-functional teams to develop RAG Pipeline and Prompt Layers, enriching contextual understanding and implementing responsible AI strategies under tight deadlines to deliver products within specified timelines.
- Migrated the product from a monolithic to horizontally scaled, distributed, scalable microservices, event-driven architecture using Kafka, zookeeper and Docker enhancing scalability and performance, which in turn supported a 50% increase in user interactions.
- Streamlined deployment processes by implementing Jenkins CI/CD pipelines to automate builds, testing, and deployments across cloud environments, resulting in a 30% reduction in downtime.
- Successfully recovered an on-premise Enterprise MongoDB production server from an unrecoverable error, ensuring minimal disruption to operations.

Big Data Consultant, Voestalpine GMBH (RBHU Analytics) – Mumbai, MH, India Jun 2023 – Aug 2023

- Spearheaded AI-driven initiatives to enhance sales forecasting accuracy and strategic decision-making.
- Employed predictive modelling techniques, integrating an ARIMA model and machine learning algorithms, to achieve precise sales forecasting and demand analysis.
- Conducted comprehensive analysis of 300,000 sales data points over 6 years, utilizing AI-driven analytics tools.
- Developed bespoke visualizations and reports leveraging AI-powered insights, including an advanced PowerBI Dashboard tailored for Voestalpine High-Performance Metals GMBH
- Achieved a 15% improvement in sales forecast accuracy and a 15% reduction in inventory costs through AI-driven strategies.

Backend Developer, Vpledge – Mumbai, MH, India Dec 2021 – Jul 2022

- Improved customer service efficiency by implementing a transformer-based chatbot, reducing response time by 20%.
- Restructured GraphQL APIs, resulting in a 30% increase in Server Query response time.
- Enhanced user engagement and accessibility by creating and hosting the company's website with Nuxt.js, Google Firebase, Google Analytics, and TailwindCSS driving a 25% increase in traffic.
- Facilitated collaboration and innovation by actively engaging with internal cross-functional teams and Meta (Ireland), successfully delivering multiple projects and initiatives.

Skills

Programming: Python, R, SQL, Numpy, Pandas, Matplotlib, SK-Learn, PyTorch, TensorFlow, Keras, NLTK, Spacy, LangChain, FastAPI, REST API, GraphQL, BASH, ~~FFX~~
LLM: OpenAI (GPT, Dall.E, Whisper), LLaMA, BERT, Mistral, Claude, Ollama
NLP: Text Analytics, Sentiment Analysis, Text Classification, Information Extraction, Named Entity Recognition
Deep Learning: Neural Network Optimisation, Deep Reinforcement Learning(DQN), RLHF, GANs
Big Data: Kafka, Zookeeper, Hadoop, HDFS, Map-Reduce, Airflow, Spark
DBMS: Weaviate, Pinecone, PostgreSQL, MySQL, MongoDB, Redis, S3, Cassandra
ML Ops: Jenkins CI-CD, Git, Docker, Nginx, AWS (EC2, LAMBDA, Bedrock), Jupyter Notebook, Linux
Soft Skills: Research, Analytical thinking, Problem-solving, Teamwork, Effective communication, German Language (A2)

Education

CDAC – PG Diploma in Big Data Analytics, Grade: Valedictorian 86.88 % 2023
Coursework: Deep Learning, Reinforcement Learning, Machine Learning, Distributed Systems, Big Data Analytics
University of Mumbai – Bachelor of Electrical Engineering 2021
Coursework: Control Systems, Signal Processing, Calculus, Statistical Analysis, Linear Algebra