

Ravi Kumar

[LinkedIn](#) | [GitHub](#) | [Medium](#)

: raviikrds@gmail.com

EXPERIENCE

CAPGEMINI

Apr 2021 - Present

Associate Consultant (Data Scientist)

Dec 2023 - Present

- Built and deployed end-to-end ML pipelines for text analytics, data ingestion, preprocessing, model training, and prediction in production environments.
- Developed NLP models for text classification using both traditional algorithms (LSTM, SVM, Naive Bayes) and LLMs, working with bilingual datasets and applying techniques such as embedding-based feature engineering (Word2Vec, BlazingText), attention mechanisms, and model fine-tuning.
- Automated deployment workflows using CI/CD and MLOps practices, enabling scalable, reproducible rollouts with integrated monitoring and model versioning.
- Delivered scalable ML solutions using AWS (SageMaker, Lambda, S3), GCP, and Databricks, selecting infrastructure based on cost-efficiency, model latency, and deployment needs.
- Collaborated cross-functionally with data engineers, product manager, and business stakeholders to align ML models with real-world use cases, streamline integration, and drive continuous improvement through feedback loops.

Senior Software Engineer (Data Scientist)

June 2022 - Dec 2023

- Designed and deployed end-to-end solutions for text classification, topic modeling, and topic extraction using Python, ML framework.
- Applied both supervised and unsupervised learning techniques, including LDA for topic modeling and clustering on unstructured text data.
- Built scalable ML pipelines covering data preprocessing, feature extraction, and model training, enabling continuous iteration and experimentation.
- Deployed models on AWS SageMaker, integrating with S3, Lambda, and other AWS services for real-time inference and automated retraining workflows.
- Collaborated with cross-functional teams and mentored junior engineers on NLP deployment best practices.

Software Engineer (Data Analyst)

Apr 2021 - June 2022

- Extracted and transformed large-scale datasets from AWS RedShift using SQL and Python Libraries to support analytics workflows and business reporting.
- Designed and maintained Power BI dashboards to visualize KPIs, user trends, and system performance for data-driven decision-making.
- Automated recurring reporting tasks using reusable ETL scripts, improving efficiency and reducing manual overhead.
- Worked closely with product managers, and cross-functional teams to define metrics, ensure data accuracy, and present insights through clear storytelling.

PROJECTS

Smart FAQ Chatbot

HuggingFace Embeddings, LLM, FastAPI, Streamlit, FAISS

- Developed an intelligent FAQ chatbot using Gemini LLM embeddings and FAISS-based semantic vector search over a knowledge base.
- Built a scalable backend with FastAPI (deployed via Render) and an interactive Streamlit frontend (Streamlit Cloud), enabling natural language Q&A.
- Implemented CI/CD with GitHub Actions for automated backend deployment and ensured secure, efficient system architecture.

NINA IT Assistant Agent

Google ADK, Vertex AI, Agentspace, LangChain, Gemini, Buckets, ITSM

- Built Nina, an AI-powered ITSM assistant that automates ticket triage, incident analysis, and knowledge base suggestions using LLMs and RAG.
- Integrated with ServiceNow via custom Python utilities for real-time incident monitoring and resolution assignment, deployed scripts and a centralized agent.
- Implemented a CI/CD pipeline with GitHub Actions and Dockerized deployment, enabling scalable and secure automation workflows for IT service operations.

Service Ticket Dashboard

SQL, Power BI, Python

- Developed an interactive dashboard to monitor service ticket trends, KPIs, and resolution metrics, empowering operational teams with data-driven insights.
- Automated data extraction and transformation from structured sources using SQL and Python, delivering accurate insights with reduced manual effort.
- Implemented dynamic filters, drill-down capabilities, and role-based access in Power BI to tailor insights for business and support stakeholders.

EDUCATION

LAKSHMI NARAIN COLLEGE OF TECHNOLOGY

Bhopal, Madhya Pradesh

- B.E. in Electronics and Communication Engineering
- CGPA: 8.19/10

June 2016 - June 2020

SKILLS

Programming Languages & Frameworks: Python, SQL, Pandas, NumPy,

Machine Learning & AI: Regression, Classification, Clustering, SVM, LSTM, LLMs, RAG, Agentic AI

Data Analytics & Statistical Modeling: EDA, Feature Engineering, A/B Testing, Hypothesis Testing

NLP & Language Models: Text Classification, Word2Vec, Blazing Text, LDA, LangChain, Prompt Engineering

Data Visualization & BI Tools: Power BI, Matplotlib, Seaborn

Cloud Platform & Deployment: AWS SageMaker, S3, Lambda, GCP, Docker, FastAPI, GitHub Actions, CI/CD

Soft Skills: Problem Solving, Data Storytelling, Team Collaboration, Leadership, Mentorship, Agile Workflow

ACCOMPLISHMENTS

- Top 9 Finalist in Google Cloud Agentic AI Hackathon 2025 for building a multi agent IT assistant using LLMs for task automation.
- Delivered a webinar on Email Classifier & Ticket Classification attended by 500+ professionals, promoting knowledge sharing across teams.
- Published multiple blogs on Medium covering Data Science, Machine Learning, & AI contributing to the community and knowledge sharing.