RISHAV KUMAR

Machine Learning Engineer

Github | LinkedIn |Portfolio | Leetcode | rishav.dgi@gmail.com | +91 9310230085 | New Delhi, India, 110030

SUMMARY

Machine Learning Engineer at Tata Consultancy Services with **3.5+ years** of IT experience in **Machine Learning, Data Analysis, GenAI**, reinforced by multiple real-world projects and proficiency in **Python, FastAPI, Sklearn, TensorFlow, Pytorch, AWS Cloud** and **SQL** with strong focus on architecting intelligent and scalable solutions.

SKILLS

- **Programming:** Python, SQL
- **Technology:** Artificial Intelligence, Machine learning, Data science, Statistics, Deep learning, NLP, GenAI (Generative AI), LLM (Large Language models), Predictive Modeling, Exploratory Data Analysis (EDA), ETL, Forecasting, Data Wrangling, Feature Engineering, Data Cleaning Preprocessing, Web Scraping, Power BI, Tableau, Docker
- **Generative AI:** LangChain, LangGraph, VectorDB, GraphDB, CrewAI, MCP, LangSmith
- **Framework, Cloud and MLOps:** Pandas, NumPy, Scikit-Learn, TensorFlow, Keras, PyTorch, Matplotlib, Seaborn, AWS (S3, EC2, Lambda, ECR, CodeDeploy, IAM), MLflow, Git, CI/CD Pipelines

EXPERIENCE

Tata Consultancy Services Ltd.   Jun 2022 – present
Machine Learning Engineer   Gurugram - India

Rabobank One Bank One Flow:
- Retain money within the bank by giving offers to existing clients and onboarding new potential corporate customers.
- Developed an end-to-end pipeline using pre-trained **Roberta NER** based **transformers** and Facebook's **Fasttext** for **intent classification** to identify customer's industry by their names.
- Utilized **Levenshtein distance** and **Rapid Fuzz** NLP algorithms for name matching to identify customer's existence in the bank, achieved **90% accuracy**.
- Created an **RFM** score model to prioritize leads into hot, medium, or cold categories.
- This use case helped Rabo Bank to generate a revenue of **7Cr** in 2025.

Rabobank Mortgage:
- Predict which customers are most likely to apply for a home loan (propensity scoring) and segment mortgage customers to improve targeted marketing strategies.
- Developed a **Decision Tree** algorithm to predict the likelihood of a customer applying for a home loan.
- Used **LIME** and **SHAP** for explainability.
- Applied **K-Means clustering** for mortgage customer **segmentation** into categories like Elite, Loyal, and Ace.

RaboBank Other Channel Movement Prevention:
- Activate debit card inactive customers who use debit cards only for cash withdrawals but not for online transactions.
- Conducted **EDA** using **Power BI** and implemented multiple **regression** and **regularization** techniques using the TPOT library.
- Predicted customer MAB (Monthly Average Balance) for the next two months to help RMs target inactive customers.

TCS Banks Application Engineer:
- Monitored and managed critical batch and job runs ensuring smooth execution and minimal downtime.
- Conducted quantitative analysis on capital market data using SQL and Python (Pandas/NumPy), generating actionable insights that optimized wealth management strategies. Automated reports building, reduced manual effort up to **90%**.
- Developed a **predictive model using a neural network** to estimate completion time of TCS BaNCS application batches and job runs, enabling proactive measures to avoid potential delays or failures in stock market settlements and high-value transactions.
- Engineered a **hybrid recommendation system** for financial products, leveraged explicit data and implicit data to enhance recommendation accuracy achieving an overall **F1-score of 87%.**

EDUCATION

A.P.J Abdul Kalam Technical University   **Aug 2018 – Jun 2022**
Bachelor of Technology in Computer Science Engineering CGPA: 8.1/10   Noida, Uttar Pradesh

PROJECTS

Advanced RAG Chatbot with Hybrid Search | (Python, LangChain, OpenAI, ChromaDB, FastAPI)
- Enhanced Retrieval System: Implemented hybrid search combining dense vector embeddings with sparse lexical search, improving retrieval accuracy by 35% over baseline RAG.
- Optimized Chunking Strategy: Developed semantic chunking with recursive text splitting and overlap handling, reducing context fragmentation by **40%**.
- Multi-format Document Support: Engineered processing pipeline for PDF, DOCX, and TXT files with automatic metadata extraction and structured storage.
- Production Deployment Containerized application using Docker and deployed with load balancing, serving 500+ daily queries with **99% uptime**.

Enterprise Resume Optimization Platform | (RAG, AWS Bedrock, Lambda, API Gateway, S3, FAISS)
- Semantic Search Engine: Built FAISS-based similarity search with cosine similarity scoring, retrieving top 5 most relevant job descriptions with **94%** accuracy.
- Serverless Architecture: Designed event-driven pipeline processing 1000+ resumes daily with auto-scaling and cost optimization.
- ATS Compliance: Integrated parsing for **20+** ATS keywords and formatting requirements, increasing client interview rates by 40%.
- Real-time Generation: Deployed Claude/Titan models via Bedrock for instant summary generation with 2-second average response time.

Crypto Signals with LSTM |**Github**
- This project aims to predict the price of Bitcoin (BTC) against the US Dollar (USD) using historical data and a Long Short-Term Memory (LSTM) neural network. The model is built using TensorFlow and Keras, and it leverages data from Yahoo Finance.

SMC Trading bot with Gemini |Github
- This is a streamlit-based AI trading assistant that analyzes uploaded chart images using Google's Gemini Vision model and provides sniper trade setups based on ICT Smart Money Concepts (SMC). The model returns Buy/Sell signals, entry, stop loss, and a detailed rationale.

University shortlisting and prediction system using Linear Regression and Streamlit | Github
- This repository contains a project focused on shortlisting universities for prospective students based on their profiles and predicting the likelihood of admission. The project leverages Linear Regression to provide recommendations and predictions.

CERTIFICATIONS & COURSES

- Agentic AI Bootcamp (LangGraph, Langchain)
- Building Gen AI App 12+ Hands-on Projects with Gemini Pro
- Complete Generative AI Course with Langchain and Huggingface
- Data Science Masters Certification Program
- The Complete Python Bootcamp from Zero to Hero in Python
- Mastering Data Structures and Algorithms

ACHIEVEMENTS

- Secured All India Rank 48 in TCS CodeVita Season 10 and received FTE offer from TCS Digital.
- Runner-up in the TCS AI Innovation Hackathon 2023, recognized for developing an innovative AI-based solution.
- Honored twice as Employee of the Quarter for exceptional performance and contribution.