Hrithik Rai Saxena

hrsaxena97@gmail.com +49-15172456126 Linkedin GitHub Portfolio medium

Work Experience - 4 years 11 months

Sagar Institute of Science, Technology and Research Bhopal, India
Assistant Professor (CSE – AI/ML) December 2024 – Present

- Designed and taught graduate-level courses on Machine Learning, Deep Learning and GenAI.
- Skill Development and mentoring project work of around 300 students.
- Conducting research based on neurocomputing - Grid and Place neuron modelling for spatial navigation.

Livello Technologies Düsseldorf, Germany
Data Scientist May 2024 – Aug 2024

- Led the development of revenue and product demand forecasting systems incorporating AI-research, statistical modelling, backend, and infrastructure.
- Developed CV on edge by implementing data transformations and backend migration pipelines for demographic analysis on edge device – Nvidia Jetson Nano.
- Managed production releases on GCP-Cloud Run using CI/CD pipelines (GitLab), ensuring compliance with IT policies.
- Reduced Cumulative Forecast Error from 8 to 3.2 through probabilistic modeling for intermittent demand (CROSTON-TSB).

Livello Technologies Düsseldorf, Germany
Research Student – Master Thesis **Nov** 2023 – May 2024

- Spearheaded dynamically fine-tuned SARIMA-X models across kiosks using multivariate Time Series Analysis, hyperparameter optimization, and Python/R for EDA and statistical analysis.
- Optimized forecast system's microservice architecture, integrating best practices and ensuring modular development.
- Resolved high response times, memory leaks, and downtimes, improving system reliability.
- Collaborated with cross-functional teams for A/B and user acceptance testing (UAT), fostering strong stakeholder relationships.
- Key achievement – Top Star Performer

Livello Technologies Düsseldorf, Germany
Data science intern **Aug** 2023 – Nov 2023

- Introduced exogenous variables (holidays, weather) and Time-GAN based data imputation reducing MAPE from 32% to 15%.
- Developed new features for the forecast service, including dashboard metrics and ranking demand tables, using advanced EDA procedures.
- Built efficient ETL pipelines (MongoDB to BigQuery), ensuring data integrity and security.
- Documented procedures and best practices, aligning with compliance standards in Confluence.

SYSKRON GmbH (Krones) Regensburg, Germany
AI Intern Oct 2022 - Mar 2023

- Formulated retraining strategies for Reinforcement Learning Product - ContiloopAI, emphasizing experiment design and monitoring.
- Enhanced data preprocessing pipelines with feature engineering for raw factory data, enabling real-time sampling strategies.
- Authored technical documentation on reinforcement learning and AI training, contributing to team knowledge base.

SCHMP, Government of India (MP) Bhopal, India
Data Scientist and Analyst (Contract) Sept 2018 – July 2021

- Deployed end-to-end SVM based ml application for classification of Ration card allotment reducing turnaround time by 70%.
- Reengineered legacy information architectures, aligning with advanced data science strategies and dashboards.

Education

Deggendorf Institute of Technology Master of Science (M.S.) - Major in Artificial Intelligence and Data Science
Deggendorf, Germany Sept 2021 – Mar 2024

University Institute of Technology, RGPV, India Bachelor of Engineering (B.E.) - Information Technology
Bhopal, India Apr 2015 – May 2019

Technical Skills

- Programming : Python, R, C, HTML, CSS, JavaScript, API development (Restful and Websockets), Hadoop, Spark
- Machine Learning and Deep Learning: Scikit-learn, Statsmodels, TensorFlow, PyTorch, AutoML, XAI (SHAP,LIME)
- Generative AI: LLM based application development and deployment, RAG based systems (memory-based multimodal), Langchain, Langserve, LoRA Fine tuning, Multilingual models, Quantized models, Vision models, Reasoning models, agentic workflows and tools (agno agents), building virtual assistants and chatbots, content generation (text-to-text, text-to-image,text-to-video), cost optimization, open source(Ollama and hugging face) models, OpenAI(GPT4o,o1), Deepseek(R1), Mistral models (quantized), Google (Gemini 2.0 flash),vector databases (Chroma, FAISS, Qdrant).
- Big Data & Databases: SQL, NoSQL (MongoDB, Redis, GraphDB), Data Lake, Data Virtualization, DB2 DWH, PowerBI
- MLOps & Model Deployment: CI/CD, Docker, MLFlow, Weights & Biases, Ray, Flyte, Terraform, Airflow
- Cloud: AWS (S3, Kinesis, Firehose, Glue, SageMaker, Lambda, Bedrock, OpenSearch), GCP (BigQuery, Cloud Run)