Yash Kumar Roy

8192090180 yashkumarroy164@gmail.com LinkedIn GitHub Portfolio

Professional Experience

Data Science Associate, Sequence Surface LLP *25 July – Present*

- Spearheaded website management and optimization with focus on SEO, performance monitoring and bug fixes, improving visibility and user engagement.
- Automated reporting and data-cleaning tasks using Excel (Pivot Tables, VLOOKUP) and Python scripts, increasing operational efficiency.
- Designed and deployed ML solutions to solve business problems using Python toolset (Pandas, scikit-learn, NumPy).

Data Science Intern, AI Variant *July 2023 – March 2024*

- Worked on data preprocessing, feature engineering and model evaluation for client projects.
- Assisted in building prototypes for NLP and predictive analytics tasks.

Projects

RAG-based Document Question Answering System using LlamaIndex *April 2025*

- Built a Retrieval-Augmented Generation pipeline using OpenAI GPT-3.5 Turbo and HuggingFace transformers for embeddings.
- Used LlamaIndex for document ingestion, chunking and vector storage; supported semantic querying over PDFs with LLM-powered answers.

Style Transfer using GAN *June 2025*

- Implemented a style-transfer GAN (inspired by CycleGAN) using TensorFlow/Keras to translate grayscale medical images.
- Preprocessed MRI data, applied adversarial, cycle-consistency and identity losses; visualized outputs and tracked loss/quality curves.

Bankruptcy Risk Prediction *March 2024*

- Developed a Random Forest classifier on financial dataset (6,000+ firms). Performed outlier handling and GridSearchCV hyperparameter tuning.
- Achieved $\approx 91\%$ accuracy and improved model performance through robust preprocessing and tuning.

Skills

Languages & Web: Python, SQL, Streamlit, EXCEL
AI: LangChain, OpenAI, Machine Learning, numpy, Pandas, GenAI, NLP, Google Gemini, HuggingFace, Embeddings, RAG, Vector DBs (pgVector, ChromaDB), Prompt Engineering
Database & Tools: MySQL, Git, Power BI, Figma

Education

Liverpool John Moores University, U.K. *Jun 2024 – Apr 2026*
Master of Science: Data Science & AI
IIIT Bangalore *Jun 2024 – Jun 2025*
Executive Post Graduation Programme: Data Science & AI
GBPIET, Uttarakhand *Aug 2019 – Jun 2023*
B.Tech in Computer Science and Engineering