

# Abhishek Kumar

Mob : +44 07407267412 | Email: [abhi.kuvmr@gmail.com](mailto:abhi.kuvmr@gmail.com) | Address: Cardiff, United Kingdom CF24 1DL

Portfolio: <https://abhishekokumar.github.io/portfolio/> | LinkedIn: <https://www.linkedin.com/in/abhi-km/>

## PROFILE

Data Scientist & Analyst with **5+ years of combined experience** across Analyst & Software Engineering experience with an MSc in Data Science & Analytics (Distinction, Cardiff University). **Expertise in Python, SQL, BI dashboards, ETL pipelines, Data Cleaning, and machine learning** and dashboard reporting. **Experienced in predictive modelling, NLP, Generative AI, chatbot Building and risk forecasting**, with a track record of extracting business-critical insights from both structured and unstructured data.

## EDUCATION

MSc Data Science & Analytics (Distinction) – Cardiff University, UK	(Sep 2023 – Dec 2024)
B.Tech Computer Science & Engineering (First Class) – AKTU, India	(Aug 2016 – Sep 2020)

## KEY SKILLS

- **Programming & Analytics:** Python (pandas, NumPy, scikit-learn, TensorFlow, PyTorch), MySQL, A/B Testing
- **ETL & Data Pipelines:** PySpark, data wrangling, validation, cleansing, job scheduling
- **Machine Learning & AI:** Regression, classification, forecasting, anomaly detection, NLP, BERT embeddings, LLMs, Hugging Face, prompt engineering, FastAPI deployment, MLflow
- **Business Intelligence & Visualisation:** Power BI, Excel (advanced), data storytelling, Reporting
- **Cloud & Tools:** AWS, Git/GitHub, Docker, CI/CD familiarity
- **Professional:** Stakeholder Engagement, agile (Scrum, Jira), cross-team collaboration

## RELEVANT EXPERIENCE

- |  |                              |
|--|------------------------------|
| <b>Data Analyst (Part-Time), Euroclick Ltd, Cardiff, UK</b>  | <b>(Mar 2024 – Present)</b>  |
| <ul style="list-style-type: none"><li>• Designed and maintained <b>Power BI dashboards</b>, boosting sales by <b>15%</b> and reducing stock inefficiencies by <b>40%</b></li><li>• Wrote and optimised <b>SQL queries</b> to extract, validate, and cleanse retail datasets</li><li>• Partnered with sales and marketing teams to define data requirements and deliver actionable insights</li></ul>                   |                              |
| <b>Software Engineer - Payments &amp; Data Processing, Kissht, Mumbai, India</b>   | <b>(Jan 2023 – Sep 2023)</b> |
| <ul style="list-style-type: none"><li>• Developed microservices for payments &amp; disbursements, <b>improving transaction processing</b> speed by <b>20%</b></li><li>• Integrated APIs from banks/PSPs (Yes Bank, RazorpayX, Cashfree) for real-time financial data handling</li><li>• Optimized MySQL-based services for <b>high-volume data pipelines</b>, enhancing reliability and compliance reporting</li></ul> |                              |
| <b>Software Developer - Ecommerce Analyst, Cedcross Technologies, Lucknow, India</b>   | <b>(Jan 2020 – Jan 2023)</b> |
| <ul style="list-style-type: none"><li>• Analysed customer/order data across <b>20+</b> e-commerce marketplaces, surfacing trends to reduce refund-related queries</li><li>• Automated reporting workflows with MySQL &amp; Python, minimizing manual tasks and improving data accuracy</li><li>• Facilitated integration of <b>20+</b> platforms while embedding analytics into operational decision-making</li></ul>  |                              |

## PROJECTS

- **Gen AI Log Classifier:** Built a pipeline with **Regex, BERT embeddings, LLaMA LLM**, and **Logistic Regression** to classify system logs. Deployed via **FastAPI**, enabling real-time anomaly detection, cut costs by **40%** and boosting resource efficiency
- **Energy Usage Prediction:** Developed **ML models (Random Forest, LGBM)** achieving **more than 90% accuracy**, supporting residential energy planning at Kaggle Dataset
- **River Flow Forecasting (Dissertation):** Forecasted daily river flows by building machine learning models (regression, Neural Network), and anomaly detection for environmental risk with **greater than 96% accuracy** by acquiring Large Dataset.
- **AI Chatbot (Dialogflow, GCP):** Created NLP-driven for Restaurant with SQL backend, **automating 95% of queries and reducing support workload by 80%**
- **E-commerce Sales Dashboard:** Designed real-time KPI dashboards on sales and customer behaviour, **cutting manual reporting time by 70% and enabling faster decision-making**

## ACHIEVEMENTS

- 5 star badges in python programming at Hacker Rank