

GLAWIN JOHN ALVA

Birmingham, United Kingdom ◊ +44 7341704465

glawin24@gmail.com ◊ [LinkedIn Profile](#) ◊ [GitHub Portfolio](#) ◊ [Personal Website](#)

SUMMARY

Data & AI Engineer with over a year of hands-on industry experience across **data engineering**, automation, and applied machine learning, backed by a Distinction in **MSc Artificial Intelligence & Business Strategy** (Aston University). Skilled in designing efficient **data pipelines** and batch/near-real-time workflows, developing and optimizing ML models, and integrating end-to-end solutions using **Python**, **SQL**, **Apache PySpark**, and **AWS cloud services**. Proven ability to collaborate with product and engineering teams to deliver **production-grade** features, improve **data quality** and **reliability**, and enable **analytics and BI** through tools like **Power BI / Looker**. Focused on **scalability**, **performance**, and measurable business impact.

TECHNICAL SKILLS

Programming: Python, SQL (PL/SQL, T-SQL), Java, C++

Cloud & Data Platforms: AWS (S3, Glue, Lambda, Athena, Redshift – fundamentals)

Data Engineering: ETL/ELT, Apache PySpark, Data Lakes, Pentaho, PostgreSQL, Greenplum

Visualization & BI: Power BI, Google Looker, Matplotlib, Seaborn

Machine Learning (Applied): Scikit-learn, TensorFlow, PyTorch

Data Practices: Data Modeling, Data Quality, Metadata Management

Soft Skills: Strategic Thinking, Collaboration, Problem Solving, Leadership

EXPERIENCE

Software Developer (Data-Focused), VeritaPort

Jun 2025 – Present

Birmingham, United Kingdom

- Contributing as a core developer to **Verigo**, a gamified navigation app encouraging eco-friendly travel through challenges and rewards.
- Designing and implementing scalable **backend systems** and **data workflows** using **PostgreSQL**, React Native, Node.js, and third-party APIs (TomTom / Google Maps).
- Building in-app **event tracking**, **data models**, and XP/badge logic to support **analytics and user behaviour insights**.
- Developing **AI-supported data pipelines** for user behaviour analysis and route optimisation.
- Integrating **CI/CD pipelines** to improve deployment reliability and data quality.
- Leading development of the flagship project — **AI-Powered Traffic Audit**, leveraging **computer vision and machine learning** for **anomaly detection** and traffic insights. ([GitHub Repository](#))

Associate IT, Tata AIG General Insurance Company Limited

Aug 2022 – Aug 2023

Mumbai, India

- Developed and maintained a scalable **Data Lake** using **Apache PySpark** and **PostgreSQL (Greenplum)**, improving data accessibility and query performance.
- Processed and cleaned over **100,000 records**, enhancing data accuracy and consistency across business systems.
- Automated recurring **ETL workflows** with **Pentaho**, reducing manual effort by 40% and increasing operational efficiency.
- Collaborated with cross-functional teams to ensure data integrity, scalability, and compliance with internal governance standards.
- Contributed to documentation and best practice frameworks for data pipeline reliability and maintenance.

KEY PROJECTS

Energy Consumption Forecasting & Anomaly Detection

Jun 2024 – Sep 2024

Aston University, United Kingdom

- Built an integrated system combining **energy forecasting and anomaly detection** using a hybrid **LSTM + Stationary Wavelet Transform (SWT)** model.
- Applied **deep learning and unsupervised algorithms** such as Isolation Forest, Autoencoders, One-Class SVM, and LOF to detect abnormal usage patterns.
- Achieved a forecasting **RMSE of 0.022 on test data**, enabling proactive energy management.
- **Tools:** Python, TensorFlow, Scikit-learn, Pandas, SWT, SMOTE ([GitHub Repository](#))

Big Data Road Safety Risk Prediction

May 2024

Aston University, United Kingdom

- Developed ML models to predict **road collision severity (Slight, Serious, Fatal)** using the UK DfT 2023 dataset.
- Trained and optimised **Random Forest** and **LightGBM** classifiers, improving accuracy from 68% to 75.4%.
- Identified key predictors such as road type, lighting conditions, and casualty count, supporting data-driven policy planning.
- **Tools:** Python, Pandas, LightGBM, Scikit-learn, Matplotlib ([GitHub Repository](#))

E-Commerce Customer Segmentation & Fraud Detection

Nov 2023

Aston University, United Kingdom

- Developed an end-to-end ML pipeline combining **clustering, fraud detection, and reinforcement learning** to uncover customer behaviour patterns.
- Achieved **99.67% accuracy**, 90% precision, and 81% recall in fraud detection using Random Forest and Gradient Boosting.
- Segmented 118,929 transactions into 4 key behavioural clusters, enhancing fraud prevention and marketing targeting.
- **Tools:** Python, Scikit-learn, Random Forest, K-Means, PCA, Reinforcement Learning ([GitHub Repository](#))

EDUCATION

MSc Artificial Intelligence and Business Strategy, Aston University

Sep 2023 – Oct 2024

Birmingham, United Kingdom — Distinction (70%).

BSc Information Technology, University of Mumbai

Jun 2018 – May 2020

Mumbai, India — Distinction (8.02/10 CGPA).

GNIIT – Cloud and Mobile Software Engineering, NIIT

Jun 2020 – Jun 2021

Mumbai, India — Graded **Excellent** (equivalent to **8–10 CGPA**).

CERTIFICATIONS

- [AWS Data Engineer – Associate](#)
- [NIIT Cloud and Mobile Software Engineering](#)