Junayd Elahi

London | Junaydelahi@outlook.com
GitHub: github.com/junayd-elahi | **LinkedIn:** www.linkedin.com/in/junayd-elahi-2029b9213

Technical Skills

- **Tools & Infrastructure:** dbt, Snowflake, Airflow, Docker, AWS S3, CloudFront, CI/CD, Git
- **Languages:** Python, SQL, Go, R
- **Data & Analysis:** Pandas, NumPy, Excel, Matplotlib, statsmodels, Power BI
- **Machine Learning:** scikit-learn, XGBoost, Random Forest, TensorFlow, PyTorch

Work Experience

Coldwell Banker, Southbank - Data Analyst
October 2023 - Present

- Built Python automation tools to replace recurring manual data-entry workflows, reducing weekly manual workload by 40%.
- Scraped, cleaned, and analysed sold-property data across 50 postcodes, processing around 2,500 listings into local price trends for client marketing and valuation analysis.
- Developed an OCR-based matching pipeline comparing live listings with 270+ units using floorplan extraction and statistical scoring, selecting the correct unit approximately 94% of the time.
- Automated a daily ETL pipeline on GitHub Actions and rebuilt its scheduling on a Cloudflare Worker after runs were firing up to 10 hours late, so it now starts on time every day.

Gnomen Property Management - Quality Assurance Placement
October 2025

- Performed exploratory and boundary testing on pre-release UI, identifying validation failures, race conditions, and edge-case defects.
- Reported defects to developers with reproduction steps and severity, then presented findings and improvement recommendations to the company owner in non-technical terms.

Projects

Healthcare ELT Data Pipeline

- Built an end-to-end ELT pipeline loading ~160K rows of synthetic patient data into a Snowflake data warehouse via Python, modelled with dbt across staging and dimensional marts (star schema, 2 fact tables, 6 dimensions).
- Wrote 39 data-quality tests enforcing referential integrity, orchestrated the pipeline with Airflow in Docker, and ran dbt build in CI on every pull request.

NHS Antibiotic Prescribing Outlier Detection

- Built a Python and DuckDb pipeline ingesting 24 months of NHSBSA prescribing and GP list-size data, with idempotent loads handling schema drift across releases.
- Modelled seasonally adjusted prescribing rates per 1,000 patients in SQL and identified 40 GP practices prescribing antibiotics at over 1.5x the national rate every month

Keystroke Dynamics to Detect Parkinson's Disease

- Engineered 51 timing-based features from keystroke logs (131 participants) and built Random Forest and XGBoost models with nested subject-wise GroupKFold validation to prevent leakage.

Pneumonia Detection from Chest X-Rays

- Built a PyTorch CNN to classify chest X-rays as normal or pneumonia, reaching 0.98 recall, benchmarked against logistic regression with analysis of false positives and overfitting.

Education

University of Greenwich - BSc Computer Science (Artificial Intelligence) – Upper Second-Class Honours (2:1), 69%
September 2023 - July 2026
Relevant modules: Machine Learning (70%), Artificial Intelligence (74%), Advanced Mathematics (76%), Natural Computing (73%), Statistics with R (86%)