

MOHAMMAD AAQUIB JAWED

Data Scientist

London, UK | +44 7721 875828 | mallickaaquib@gmail.com | github.com/iAaquibjawed | LinkedIn | heyengineers.com

SUMMARY

Data scientist with 5+ years of professional engineering experience and an MSc in Artificial Intelligence (Distinction — 99/100 in Big Data Analytics, top of cohort). Hands-on with PySpark pipelines over multi-million-row datasets, statistical and deep-learning modelling, and SQL-heavy production systems — with a software-engineering foundation that makes analyses reproducible and deployable. Currently teach Big Data and Machine Learning labs at the University of East London.

TECHNICAL SKILLS

Data & Analytics: Python (Pandas, NumPy), SQL (PostgreSQL, MySQL, Spark SQL), PySpark, Hive, Hadoop, Elasticsearch, large-scale data pipelines

Machine Learning: scikit-learn, RandomForest, PyTorch, TensorFlow/Keras, model evaluation & calibration, class-imbalance methods (focal loss, weighting), transfer learning

Deployment & Tooling: Docker, Git, CI/CD, Flask APIs, AWS, model registry & drift monitoring

DATA SCIENCE PROJECTS & RESEARCH

Big Data Analytics Pipeline | MSc Big Data Analytics — 99/100, top of cohort 2025

- Engineered an end-to-end PySpark + Hive pipeline over a multi-million-row HuggingFace dataset, using Spark SQL and DataFrames for large-scale processing, analysis and prediction.

Skin-Lesion Classification (medical imaging) | Research, University of East London 2025

- Rebuilt a stalled skin-cancer classification project, raising recall from a 49% baseline to 77.7% with 0.86 AUC over 27,000+ ISIC dermoscopy images — NASNetMobile transfer learning, Keras Tuner architecture search, focal loss, class-weighted training and stratified augmentation.

Pharmacy & Healthcare Decision-Support System | Sole author — research project 2025

- Designed a clinical decision-support system to NHS-scale requirements: RandomForest models for drug-interaction, demand and disease prediction, with a model registry, drift monitoring and audit logging.

Face-Recognition Evaluation Study | MSc Dissertation — Distinction (92) 2025

- Calibrated and evaluated a recognition model at a 0.45 cosine-similarity threshold over a 300-sample test set; the resulting multi-signal liveness pipeline blocked 99.3% of presentation attacks.

Plant-Disease Image Classifier | MSc AI & Machine Vision — Distinction (92) 2025

- Trained an EfficientNet-B0 classifier with transfer learning and data augmentation, systematically tuning learning rate, batch size and epochs to maximise test-set accuracy.

Mowqit | Android App, Google Play 2025

- Built and shipped a privacy-first offline Android app (React Native, TypeScript, Zustand) with Doze-proof scheduling and a zero-data-collection design (no INTERNET permission).

EXPERIENCE

Lab Tutor, Machine Learning & Big Data | University of East London 2025 - Present

- Teach postgraduate Big Data and Machine Learning lab modules — PySpark, Spark SQL, model building and evaluation — mentoring students through large-scale data analytics.

Founder & Lead Engineer | heyengineers.com — London, UK Mar 2025 - Present

- Built and run a content platform with 500+ users and ~5,000 monthly visits; designed its ML-driven content discovery and use platform data to steer growth.

Senior Software Engineer | Norstella (remote) Jul 2024 - Feb 2025

- Raised PatientMatch alert accuracy by 25% through data-driven tuning of matching logic on clinical-trial data.
- Cut study page load times 55%+ via backend and caching optimisation on high-traffic pharma-intelligence products.

Software Engineer | PropertyPistol Nov 2022 - Jun 2024

- Lifted search result accuracy 35% and cut response time 50% by analysing query patterns and tuning Elasticsearch clustering and Query DSL.

Software Developer | Edcast Jul 2021 - Oct 2022

- Optimised services and database queries, reducing API response time 30% and query execution time 50%.

Earlier: Software Engineer, Cypher IT Solutions (Sep 2019 - May 2021) — designed database schemas and data-driven MVC pages, cutting data-retrieval time 30%. Cyber Security Intern, Gurugram Police (2021) — identified and mitigated SQL-injection and network-security vulnerabilities.

EDUCATION

MSc Artificial Intelligence — Distinction (93% average, top of cohort) | University of East London 2024 - 2026

Big Data Analytics 99, Dissertation 92, every module 90+.

B.Tech, Computer Science & Engineering | Accurate Institute of Management & Technology 2015 - 2019

AWARDS & OPEN SOURCE

AI Hackathon — 1st Prize · Gold Star Award, Norstella · Rising Star Award, Edcast · Open source: Tor Project, GitLab · 70+ public repos, 1,000+ GitHub contributions/yr (Pull Shark x3)