

Hamza Gabajiwala

| +353-831446637 | hamzaig16@gmail.com | [GitHub](#) | [LinkedIn](#) | [Portfolio](#) | Dublin, Ireland | Stamp-4 |

SKILLS

Data Engineering	Apache Spark (batch/streaming), PySpark, Apache Flink, Apache Airflow (MWAA), OpenSearch
Cloud & DevOps	AWS (EMR, S3, Glue, Lambda, EC2), Docker, Kubernetes, Jenkins, CI/CD
Programming Languages	Python, SQL, Scala, Java
Databases & Web	MySQL, PostgreSQL, MongoDB, React, Flask, Django, Spring Boot, RESTful APIs

EXPERIENCE

Yahoo — *Software Developer I* | Dublin, Ireland **Feb 2024 - present**

- Built and owned 10+ Airflow DAGs orchestrating Spark/PySpark scoring pipelines on EMR, processing 1.28B+ events/day across 400+ segments, powering ad targeting for hundreds of millions of users.
- Implemented multi-level real-time ad targeting in a Flink streaming app by extending the Java segment cache with new index structures; enabled package/order-level matching with sub-second latency.
- Shipped Standard Audiences globally by developing the OpenSearch rules pipeline, improving 400+ segment definitions and expanding audience reach by 60% using new contextual signals.
- Migrated 4 major scoring systems to new AWS infrastructure, achieving 13% cost reduction via instance benchmarking (m6 vs c6) and replacing Glue Catalog dependencies with direct S3 reads.
- Designed and delivered a fault-tolerant reprocessing engine covering 113K+ audience segments, eliminating client-reported zero-count incidents through automated health-check DAGs and Slack alerting.
- Built geo-aware audience segmentation with country-level targeting and authored the design for impression-based scoring over 554GB/day of impression data using OpenSearch + windowed aggregation.

TIAA GBS — *Software Developer* | Mumbai, India **Jun 2021 - Jun 2022**

- Re-engineered the UI automation framework by migrating 150+ test cases from Selenium to WebDriver in 14 days, integrating with Jenkins to enable fully automated nightly test runs.
- Built REST-based data ingestion flows replacing manual data pulls, reducing data-collection downtime by 50% and enabling same-day reporting for internal analytics teams.

EDUCATION

Trinity College Dublin **Sep 2022 - Aug 2023**

- **Grade: 1:1** - M.Sc. in Computer Science (Specialising in Augmented and Virtual Reality)

NMIMS University **Jun 2018 - Apr 2022**

- **GPA: 3.45/4** - B. Tech in Computer Engineering

PROJECTS

Mitigating Data Collection Challenges using cGANs - Dissertation

- Built a Conditional GAN pipeline generating synthetic images to address dataset imbalance; trained ResNet50 & MobileNetV3 classifiers (TensorFlow/Keras), improving model performance on unseen classes by 40%.

Holoway - Virtual Meeting Room

- Developed cross-platform (Oculus + desktop) VR meeting environment in Unity/C#, using TDD and XP practices. Implemented real-time multiuser networking and environment synchronization.

PUBLICATIONS

- [Comparative Analysis between InceptionResnetV2 and InceptionV3 for Attention Based Image Captioning \(IEEE\)](#)
- [Comprehensive Review of Various Optimization Algorithms for Image Captioning \(IEEE\)](#)