

Anmol Dadlani

Phone: 07503919432

Github: adadlani1

Email: me@anmoldadlani.com

Data Engineer with 5+ years across the full data lifecycle — from production web apps to ML infrastructure. MSc AI from King's College London. Specialise in orchestration (Airflow), ML platforms (SageMaker, MLflow), and cost-efficient cloud architecture, with a track record of cutting infrastructure costs and accelerating training pipelines.

Work Experience

- **Realitymine Ltd - Data Engineer** - June 2025 to Present
 - Architected and deployed AWS Managed Airflow pipelines to orchestrate AI agent-driven synthetic data collection across iOS and Android devices.
 - Owned the migration from legacy Azkaban orchestration to Airflow, modernising the team's scheduling and dependency management stack.
 - Implemented MLflow to standardise experiment tracking, model versioning, and reproducibility across ML workflows.
 - Designed and led internal AI agent workshops, upskilling colleagues on agent-assisted data quality monitoring and cutting false positive alerts by 82%.
 - Re-architected ML training infrastructure on AWS SageMaker, achieving 12x training throughput at 1.5x compute cost (24 hours to 2 hours).
- **Codeweavers Ltd/Cox Automotive Inc - Data Engineer** - January 2023 to May 2025
 - Spearheaded the rebuild of the data pipeline for live event tracking using Snowplow across automotive finance lender platforms (Black Horse, BNP Paribas, and other major car finance lenders), enhancing data collection and regulatory reporting capabilities for FCA-regulated lending journeys.
 - Implemented dbt for data transformation across lender-specific data models, streamlining processing of high-volume financial transaction data and ensuring accuracy required for downstream risk, pricing, and compliance reporting.
 - Leveraged Metabase for creating insightful visualisations, enabling data-driven decision-making within the organisation.
 - Demonstrated effective communication and leadership skills by conducting numerous internal talks and workshops, sharing insights on project undertakings and providing training to software developers.
 - Innovated the platform by building custom materialisations to incorporate Apache Iceberg tables into GCP, reducing query costs by 65% — critical for the high-volume calculation datasets demanded by tier-1 financial institutions such as BNP Paribas and Lloyds Banking Group (Black Horse).
- **Codeweavers Ltd/Cox Automotive Inc - Software Engineer** - July 2020 to January 2023
 - Joined as a placement student and grew into a full-stack developer (Angular, C# .NET, PostgreSQL, AWS) delivering features into customer-facing quotation and loan origination platforms for UK car finance lenders, while mentoring new joiners in code quality, security, and auditability practices suited to regulated finance.

Education & Qualifications

- **SAF Approved during tenure at Codeweavers Ltd**
- **King's College London** - (September 2022 to August 2023)
 - **MSc Artificial Intelligence** - Pass with Merit
Key Modules covered:
Agents and Multi Agent Systems; Computer Vision - Matlab; AI Reasoning and Decision Making; AI Planning; Data Mining; Deep Learning.
- **Loughborough University** – (Oct 2018 to August 2022)
 - **BSc Computer Science and Artificial Intelligence with a Diploma in Professional Studies** – First

Class Honours

Key Modules covered:

Mathematics for Computer Science; Object Oriented Programming – Java; Databases – SQL; Advanced Networking; Human Computer Interaction.

- **Sale Grammar School Sixth Form** – Dates attended (Sep 2016 to Aug 2018)
A Levels: Chemistry, Biology, Mathematics

Projects

- **Social Bias in Large Language Models - MSc Project - Achieved 82%** - January 2023 to August 2023

Three different sub projects were involved using:

- Explore prevalence of stereotypical biases in image captioning datasets
- Determine alignment between demographic groups and concepts in datasets
- Analyse captions and associated images in Flickr30k dataset
- Focus on 'target terms' related to profession from StereoSet
- Goal is to understand the presence and correlation of biases in the dataset

Technologies used: Python3, Numpy, Pandas, NLTK, Sci-Kit Learn, Huggingface.

- **Data Mining Project - Module Coursework MSc Module - Achieved 88%** - March 2023

Three different sub projects were involved using:

- Decision trees with categorical attributes - Involved cleansing the data and building decision trees to classify the categories
- Cluster Analysis - Using Kmeans and Agglomerative clustering to find relationships between wholesale products that could be bought together.
- Mining Text Data - Involved using Natural Language Processing of Coronavirus Tweets to be able to predict the consensus of a tweet being Positive, Neutral, or Negative.

Technologies used: Python3, Numpy, Pandas, NLTK, Sci-Kit Learn.

- **Error Detection on Websites - Final Year BSc Project - Achieved 82%** - September 2021 to May 2022

- Aim is to alert Codeweavers' developers about code changes that may cause production failure
- Data sourcing using a custom web-scraping bot hosted on a Raspberry Pi and saved on GCP for ease of access
- Supervised learning was used, with a program cycling through website screenshots and users confirming functionality
- A Machine Learning model is created to predict website failure based on inputs
- A website link is given as input by the user and a screenshot is automatically taken with a performance score being generated
- The model is hosted on Heroku using Flask for API access

Technologies used include Python, Flask, Jupyter Notebook, Angular, Tensorflow, Selenium, Google Cloud, Raspberry Pi, Computer Vision, and Data Mining.

Key Skills

GCP	Cloud Scheduler	Cloud Functions	dbt	Tensorflow
Python	Apache Iceberg	EMR Serverless	PowerBI	Jupyter
SQL	PostgreSQL	BigQuery	Airflow	SageMaker AI
AWS	C#	Athena	MLFlow	Machine Learning

References

Available on request.