

NAITIK MANDALIA

Data Engineer

Melbourne, VIC, Australia • +61 459 502 944 • naitikdm@gmail.com

linkedin.com/in/naitik-mandalia-1ba7a720b • naitikmandalia.com

Full working rights in Australia — Temporary Graduate visa (subclass 485), valid to 19 August 2028

PROFESSIONAL SUMMARY

Data Engineer with 2+ years delivering production data pipelines for client projects as an independent contractor, following two years in front-end and 3D web development. Master of Data Science (Professional) from RMIT University, merit scholarship recipient. Hands-on with Python, SQL, Apache Airflow, dbt, and AWS, building ETL/ELT pipelines and data-quality checks, with foundational hands-on experience in Retrieval-Augmented Generation using the Hugging Face and OpenAI APIs. Combines statistical foundations in Bayesian modelling and hypothesis testing with software-engineering discipline to ship reliable, cost-efficient data products. Full working rights in Australia; no sponsorship required.

TECHNICAL SKILLS

Programming & Scripting: Python, SQL, R, Bash

Data Engineering & Pipelines: ETL/ELT design, Apache Airflow (orchestration), dbt (modular transformations), batch ingestion, dimensional data modelling (star / snowflake schemas), data warehousing, REST API integration, data quality and validation, CI/CD

Machine Learning & AI: Scikit-Learn, TensorFlow, Keras, CNNs, NLP, feature engineering, Bayesian modelling, A/B and hypothesis testing, model evaluation and validation

Generative AI & LLMs: Hugging Face Transformers, OpenAI API, Retrieval-Augmented Generation (RAG), vector embeddings and semantic search, prompt engineering (working knowledge)

Cloud & Databases: AWS (S3, Lambda, EC2, DynamoDB, API Gateway), serverless services; Microsoft SQL Server, DynamoDB

Visualisation & Tools: Power BI, matplotlib, seaborn, Advanced Excel, Git, GitHub, JIRA, Agile delivery

PROFESSIONAL EXPERIENCE

Naitik Mandalia (Sole Trader) — Melbourne, VIC, Australia

Freelance Data Engineer | June 2024 – Present

- Delivered end-to-end data-engineering projects for 4 small-business clients, from requirements gathering through deployment and handover.
- Built and maintained ETL/ELT pipelines in Python and SQL, ingesting data from REST APIs, relational databases, and flat files into cloud storage and warehouses.
- Orchestrated scheduled workflows in Apache Airflow, implementing DAG dependency management, retries, and failure alerting to keep client pipelines running unattended.
- Developed modular transformation models in dbt with version control and automated testing, keeping client reporting reproducible and auditable.
- Deployed pipelines and services on AWS (S3, Lambda, EC2, DynamoDB, API Gateway) using serverless patterns to keep client running costs low.
- Built Power BI dashboards and automated reports, replacing manual spreadsheet workflows for client stakeholders.
- Contributed to a Retrieval-Augmented Generation application alongside senior data engineers, working hands-on with the Hugging Face and OpenAI APIs, vector embeddings, and semantic search over client documents.

CURU App — Melbourne, VIC, Australia

Data Analytics Intern | July 2024 – October 2024

- Extracted, cleaned, and analysed large customer-feedback datasets with SQL and Python, applying feature engineering and statistical analysis to surface insights that informed product features and user engagement.
- Automated ETL collection and preprocessing pipelines in Python, eliminating recurring manual effort across analytical workflows.
- Delivered Power BI and Python visualisations and stakeholder reports that enabled data-driven decisions across product, marketing, and commercial teams.
- Executed A/B and hypothesis tests to validate the impact of new features on engagement and satisfaction, communicating findings clearly to senior stakeholders.
- Identified and resolved data-quality issues and anomalies across operational datasets, maintaining data integrity across all reporting outputs.

Ark Web — Gujarat, India

Web Designer & Developer | July 2020 – August 2022

- Developed responsive websites and interactive 3D web experiences using Three.js, WebGL, HTML5, CSS3, and JavaScript.

- Translated Figma and Photoshop designs into pixel-accurate, mobile-first interfaces, tuning rendering and asset delivery for front-end performance.
- Delivered client projects end to end, from design handoff through deployment, testing, and ongoing maintenance.

KEY PROJECTS

ChatGuardian — AI Bias Detection Tool | *Python, NLP, Hugging Face Transformers, JavaScript*

- Engineered a real-time monitoring system that detects and flags bias in AI-generated text, combining automated classification pipelines, user alerts, harm categorisation for compliance, and a continuous model-improvement feedback loop.

Colon Cancer Histology Classification | *Python, TensorFlow, Keras, Scikit-Learn, CNNs*

- Achieved 92.0% binary and 85.7% multiclass accuracy on unseen data by benchmarking logistic regression, SVM, random forest, and CNN architectures under patient-level stratified validation, with duplicate detection and class-imbalance analysis to mitigate data-leakage risk.

Online Music Subscription Platform | *AWS EC2, S3, DynamoDB, Lambda, API Gateway, Python*

- Designed and deployed a cloud-native web application implementing user authentication, subscription management, serverless Lambda processing, and REST API integration via API Gateway, hosted on an Apache server on an Ubuntu EC2 instance.

Bayesian Regression — Property Price Prediction | *R, JAGS, MCMC*

- Produced 95% credible-interval price estimates by implementing a Bayesian linear regression model that combined expert-informed priors with Melbourne market data, validated through multi-chain MCMC sampling and convergence diagnostics.

EDUCATION

Master of Data Science (Professional) | August 2022 – June 2025

RMIT University — Melbourne, VIC, Australia

- Recipient of a 20% International Merit-Based Scholarship, awarded on academic merit across a competitive postgraduate cohort.
- Relevant coursework: Computational Machine Learning, Applied Bayesian Statistics, Data Mining, Cloud Computing, Practical Data Science with Python, Advanced Programming for Data Science, Data Wrangling, Database Concepts, Algorithms & Analysis, Applied Analytics, Data Visualisation & Communication, Case Studies in Data Science, The Data Science Professional, Programming Fundamentals.

Bachelor of Computer Engineering | 2018 – 2022

Darshan Institute of Engineering & Technology — Gujarat, India

- Graduated with 85.5%. Focused on software development, object-oriented design, agile project management, and networking.

References available on request.