Eashwar Subramanian

Data Scientist | Machine Learning & Analytics | Cloud Data Systems

Melbourne, Australia | eashwars2001@gmail.com | +61 450 332 782 | **LinkedIn** | **GitHub**

SUMMARY

Early-career Data Scientist with hands-on experience designing and delivering end-to-end data, analytics, and machine learning systems. Skilled in collecting, cleaning, and transforming structured and unstructured data to generate actionable insights and support data-driven decision-making. Experienced in applying statistical and machine learning techniques, building reproducible analytical workflows, and translating technical outputs into decision-ready insights for cross-functional stakeholders. Known for a rigorous analytical approach, strong collaboration, and delivering reliable, high-quality analytical outputs.

PROFESSIONAL EXPERIENCE

Data Science Intern — Jan 2026 – Present
Culture Infusion — Melbourne, Australia

- Supporting data analysis and applied ML initiatives in the DEI and cultural analytics domain, demonstrating a committed approach to careful data preparation, exploratory analysis, and insight generation.
- Contributing to the design and implementation of data pipelines for monitoring public signals related to DEI activity, including news articles, public statements, and online discourse.
- Assisting with text-based analysis workflows while working flexibly across evolving data sources and research requirements to enable comparison of sentiment and thematic patterns across time periods.
- Supporting early-stage analytics and data product concepts by structuring unstructured text data into analyzable datasets for internal tools.

Data Science Intern — July 2025 – Nov 2025
Solara Health — Melbourne, Australia

- Built the full ingestion and semantic chunking pipeline (PDF/HTML parsing, MiniLM embeddings, metadata governance, deduplication via SHA-256), enabling clinically reliable document retrieval across 3 data tiers.
- Designed tenant-aware retrieval logic using hierarchical routing (hospital → national → state) and PostgreSQL RLS, ensuring TGA-compliant information segregation across multi-hospital deployments.
- Implemented the end-to-end evaluation framework (RAGAS) generating quantifiable metrics for faithfulness (69.4%), answer relevancy (54.3%), context precision (92.7%), and recall (91.2%).
- Performed exploratory analysis on document distributions, metadata patterns, and retrieval behaviors, providing insights that guided chunking thresholds, embedding selection, and ingestion priorities.
- Led Sprint 4 delivery, resolving merge conflicts, coordinating Jira tasks, and unblocking the pipeline for cross-functional contributors.

Data Analyst Intern — Dec 2023 – Feb 2024
PrepInsta Pvt Ltd (Remote) — India

- Delivered 5+ stakeholder dashboards (Tableau, Excel) increasing client engagement by 10%.
- Optimized SQL query performance reducing data retrieval time by 25% for executive reporting.
- Automated reporting workflows with Python (Beautiful Soup) cutting manual effort by 30%.

PROJECTS

Solara Healthcare RAG System | July 2025 – Nov 2025

- Delivered a production-grade RAG system with semantic chunking, multi-tenant routing, ingestion governance, and hallucination controls, enabling clinically safe AI responses for Australian hospitals.
- Built a reproducible evaluation pipeline using RAGAS + semantic similarity proofs, producing measurable benchmarks (faithfulness 69.4%, context precision 92.7%, recall 91.2%) for ongoing model and retrieval improvement.

Australian Retail Customer Segmentation | June 2025

- Identified a high-value customer segment representing 15% of total revenue through K-Means clustering, supporting targeted marketing recommendations.
- Corrected 72.5% invalid order IDs in 5,000-transaction dataset enabling accurate RFM analysis

Climate Policy Forecasting Dashboard | July – Sept 2024

- Built SARIMA forecasting system achieving 1.2°C MAE for Australian government climate decisions.
- Integrated Flask web app with interactive Folium mapping for urban planning stakeholder support.

TECHNICAL EXPERTISE

Languages: Python, SQL, R
Cloud/DevOps: AWS (Bedrock, Aurora, S3, Lambda, EC2), Docker, CI/CD, Git
Machine Learning/ Analytics: RAG , Semantic Chunking (LlamaIndex), Sentence Transformers (MiniLM), RAGAS Evaluation, Scikit-learn, SARIMA, K-Means, Classification Models
Databases: PostgreSQL (pgvector, RLS, IVFFlat Indexing), MySQL, DynamoDB, SQL Server
Visualization: Power BI, Tableau, Matplotlib
Frameworks / Libraries: Fast API, Flask, Pytest, Pandas, NumPy, pdfplumber
Other Specialisations: Multi-tenant architectures, retrieval evaluation, metadata governance, PII-aware processing, safety guardrails, prompt engineering

EDUCATION

Master of Data Science
Royal Melbourne Institute of Technology (RMIT)
GPA: 3.6/4.0 | Dec 2025 | Melbourne, Australia

Bachelor of Electronics & Communication Engineering
Rajagiri School of Engineering and Technology
First Class Distinction (8.8/10) | Aug 2023 | Kerala, India

PROFESSIONAL DEVELOPMENT

Data Science Melbourne Meetup - Active Member — June 2024 – Present
Networking with 50+ analytics professionals | ML trends, Power BI, Python applications

In Progress: Microsoft Azure Data Fundamentals (DP-900) • Azure AI Fundamentals (AI-900)

Relevant Coursework: Machine Learning • Big Data Processing • Cloud Computing • Advanced Programming • Computer Vision • Probability & Random Processes