

# Eashwar Subramanian

Data Science Intern | AI Systems & RAG | Healthcare & Cloud Applications

Subclass 500 | eashwars2001@gmail.com | +61 450 332 782 | LinkedIn: linkedin.com/in/eashwar-s-02461a1b3 | GitHub: github.com/Eashwar-Subramanian

## PROFESSIONAL EXPERIENCE

### Data Science Intern

Solara Health

July 2025 – Present

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 + semantic similarity proofs) generating quantifiable metrics for faithfulness (69.4%), answer relevancy (54.3%), context precision (92.7%), and recall (91.2%).
- Developed audit-ready citation verification with metadata-driven date extraction, HTML-first links, and deterministic fallback rules, improving output traceability and reducing hallucination risk.
- Diagnosed and improved retrieval quality through embedding validation, cross-encoder reranking tests, chunk coherence audits, and multi-tenant consistency checks.
- Led Sprint 4 delivery, resolving merge conflicts, coordinating Jira tasks, and unblocking the pipeline for cross-functional contributors.

### Data Analyst Intern

PreInsta Pvt Ltd (Remote)

Dec 2023 – Feb 2024

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%

## STRATEGIC PROJECTS

### Solara Healthcare RAG System | June 2025 July – 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 high-value customer segment representing 15% of total revenue through K-Means clustering
- 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

**ML / Analytics:** RAG (Retrieval-Augmented Generation), 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, Pedantic, Pandas, NumPy, BeautifulSoup4, pdfplumber

**Other Specialisations:** Multi-Tenant Architecture, Safety Guardrails (AWS Bedrock), PII Redaction, Metadata Tagging

## EDUCATION

### Master of Data Science

Royal Melbourne Institute of Technology (RMIT)

GPA: 3.4/4.0 | Expected 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