

Hari Sreedeth

Applied Data Scientist | ML Engineering & Deployment
Randwick, NSW, 2031 | 0460 058 378 | hariisreedeth@gmail.com
LinkedIn: <https://www.linkedin.com/in/harisreedeth/> | GitHub: <https://github.com/hsreedeth>
Portfolio: <https://hsreedeth.github.io/portfolio/>

SUMMARY

Applied Data Scientist with more than 3 years of experience across health data, clinical research, and machine learning. Proficient in Python, SQL, SAS, R, and Git, with experience building reproducible pipelines, analysing large datasets, and delivering practical insights.

KEY SKILLS

- Programming language and databases: Python, R, SQL, SAS, bash/shell scripting
- Machine learning and engineering pipelines: Supervised, unsupervised, deep learning, end-to-end pipelining, feature engineering, proficient in libraries such as Keras, TensorFlow, PyTorch and sklearn
- Tools and frameworks: Jupyter environments, workflow managers (Snakeflow and Nextflow), cloud platforms (GCP/AWS)
- Infrastructure and research-computing: Version control with Git and GitHub, Unix/Linux command line proficiency, proficient in executing and monitoring SLURM-based HPC and remote clusters.
- Communication: Strong ability to translate complex statistical outputs into actionable insights for diverse stakeholders.
- Agile/Team Fit: Collaborative mindset with a proven history of working cross-functionally to solve technical and domain-specific problems.

CAREER HISTORY

Visiting Postgraduate Researcher (Data Science Workflows) | The George Institute for Global Health
July 2025 - Present

Supporting health data science research by building reproducible data workflows for large-scale cohort analysis, statistical modelling, and publication-oriented reporting.

Key responsibilities

- Build Python/R/SQL workflows for data cleaning, harmonisation, feature engineering, QC, and large-scale health data analysis.
- Translate statistical and machine learning outputs into clear tables, visualisations, reports, and manuscript-ready results.

Achievement

- Developed one of the first multi-source UK Biobank harmonisation workflows for sex-specific multimorbidity analysis across 500,000+ participants. Manuscript in preparation.

EMR Migration Data Analyst & Pharmacist | GMG Healthcare, Dubai, United Arab Emirates
January 2020 - October 2023

Supported the rollout of a new central pharmacy CRM/EMR system across 60+ retail pharmacies while continuing frontline pharmacist duties. Worked with IT and pharmacy operations to migrate medication-related records, validate data quality, and track implementation progress.

Key responsibilities

- Mapped and migrated medication, pharmacy, and operational records into the new central CRM/EMR system across multiple pharmacy sites.
- Performed QA/QC checks, tracked rollout progress, identified data issues, and reported implementation status to pharmacy and IT teams.

Achievement

- Contributed to the successful UAE-wide rollout of the new pharmacy CRM/EMR system across 60+ retail pharmacies within 3 months, enabling smoother business continuity and faster operational adoption.

RESEARCH & PROJECTS

UK Biobank: Large-Scale ML Pipeline | Master's Thesis Project | The George Institute / UNSW

- High-Volume Data Processing: Engineered a scalable Python and SQL data pipeline to extract, clean, and model complex records for a massive cohort of over 500,000 participants.
- Infrastructure & Version Control: Executed resource-intensive machine learning and statistical workflows on remote SLURM-based HPC clusters. Strictly utilised Git version control to guarantee a fully reproducible and auditable codebase for production handover.

Intelligent Voicemail System for Clinical Triage | Heidi Health, Forward Deployed Engineering Challenge

- Built and deployed a working Next.js/React/TypeScript prototype that converts after-hours clinic voicemails into a structured triage workflow with inbox, detail review, urgency flags, and backlog clustering.
- Implemented a deterministic transcript-structuring layer to extract intent, urgency, callback details, confidence scores, and review flags, preserving auditability and human clinical oversight.

Demo: <https://intellvmsysh.vercel.app/>

Walkthrough video presentation: <https://youtu.be/CCTMh1nfRQc>

EDUCATION

University of New South Wales, Sydney

Master of Science in Health Data Science (Extension), 2024 – 2025 • Graduated with Distinction average.

Gulf Medical University, UAE

Doctor of Pharmacy, 2014 – 2019 • Graduated with 3.49 GPA.

REFERENCES

Available upon request.