

Luan Tran

✉ tdtruongluan1202@gmail.com · ☎ 0423538968 · Sydney NSW

LinkedIn: louistran1202

GitHub: truongluan1202

Experience

- **Data Engineer Assistant** Sydney, Australia
05/2025 - Present
Western Sydney University (Prof. Anupama Ginige)
 - Built **ETL** pipelines (**dbt**, **Airflow**, **Snowflake**) for healthcare robotics data
 - Automated multi-source ingestion & transformation, cutting latency **45%**
 - Added **dbt** tests & monitoring, reducing manual QA **60%**
- **Data Science Intern** Sydney, Australia
05/2024 - 03/2025
Skwill AI
 - Automated **Airflow** → **Snowflake** pipelines, reduced manual data pulls **50%**
 - Built **Power BI** dashboards improving analytics efficiency **30%**
 - Led **GPT-4o** feature rollout, boosted model accuracy **12%**
- **Research Assistant** Sydney, Australia
01/2025 - 05/2025
S P Jain School of Global Management (Prof. Bilal Bahaa Zaidan)
 - Engineer **ML data pipeline** for fall-risk detection, **RMSE 2.35**, improved accuracy **15%**,
 - Applied **SHAP** & **k-fold validation** for explainable performance.
- **Co-Founder** Sydney, Australia
08/2024
CapstonBara
 - Built AI-driven mental-health platform with **real-time data pipelines** & dashboards.
 - Streamlined **backend schema** & **APIs**, improving data flow **40%**

Education

- Western Sydney University** Sydney, Australia
Master by Research (Computer Vision & Data Processing)
- S P Jain School of Global Management** Sydney, Australia
Bachelors of Data Science — Dean's List for 4 consecutive semesters

Graduated, May 2025

Certificates

- Databricks **Lakehouse Fundamentals**, **AWS Certificated Cloud Practitioner**, **Power BI Data Analyst Associate**(Microsoft Certified), **SQL (Advanced)** (HackerRank), **Data Engineer Certificate** (DataCamp), Top 100 **Shopee Code League Finals** (2022)

Technical Skills

- **Data Engineering & Analytics:** Python, SQL, R, NoSQL, Power BI, Tableau, Databricks, Airflow, Snowflake, PostgreSQL, Spark SQL, MongoDB
- **Tools & Technologies:** AWS, GCP, Docker, CI/CD, Git, Scikit-learn, Langchain, Hugging Face, MLflow, PyTorch.

Personal Projects

- **Calendara - AI-Powered Calendar Assistant** (https://ai-powered-calendar-assistant.vercel.app/)
FastAPI, Next.js, PostgreSQL, Google Calendar API, Web Search
 - Developed full-stack conversational assistant for natural-language scheduling & event retrieval.
 - Integrated **Google Calendar API** and **web search** for contextual event management.
 - Deployed multi-provider backend (**OpenAI/Gemini**) enabling smart intent routing.
- **Airtable Clone** (https://airtable-clone-eta.vercel.app/)
Next.js, TRPC, TanStack Table, PostgreSQL, Prisma
 - Built Airtable-style data grid with **virtualized** infinite scroll supporting **100k+** rows.
 - Implemented dynamic columns, filtering, and real-time collaboration via **TRPC**.
 - Optimized **PostgreSQL** queries and **Prisma ORM** for **sub-100 ms** interactions.
- **Sports Analytics Data Pipeline** (https://github.com/truongluan1202/Sports-Analytics-NoSQL-Database.git)
MongoDB, PostgreSQL, Python
 - Engineered ETL pipelines processing **1.9K+** shot events from **6** European football leagues.
 - Modeled both NoSQL (**MongoDB**) and relational (**PostgreSQL**) schemas for analytical workloads.
 - Added **triggers**, validation rules, and **indexing**, improving query speed **35%** and data reliability **40%**.