YEO MING JIE, JONATHAN

Email: yeomj@u.nus.edu | Mobile: +65 92356385 | LinkedIn | GitHub

PROFESSIONAL SUMMARY

Data Scientist and former policy analyst with two years of experience, skilled in merging policy and data modelling to drive strategic decisions and boost operational efficiency. Strong foundations in machine learning and statistical analysis. Effective communicator with experience in stakeholder engagement and cross-functional collaboration. Passionate about leveraging data to serve the public and enhance outcomes in healthcare and transportation sectors.

WORK EXPERIENCE

Ng Teng Fong General Hospital (National University Health System) **Sep 2023 - Present**

Data Analyst, Quality Improvement & Innovation

- Technical lead data scientist in the Digital Think Tank Analytics team, delivering proof-of-concept applications for clinical use cases leveraging Large and Small Language Models (LLMs/SLMs).
- Directed application development and deployment on Healthcare Commercial Cloud (AWS) and on-premises Endeavour AI (EAI), ensuring scalability, compliance, and integration with hospital systems.
- Led cross-institutional AI/ML initiatives, engaging senior management and clinicians, and contributing to cluster-level strategy discussions at the LLM Project Management Office.
- Implemented an LLM-based automation solution to streamline Emergency Department (ED) audits, scaled across 3 hospitals; achieved 100+ man-hours saved annually and doubled audit coverage.
- Developed predictive models including Neurological Deterioration Risk and Length-of-Stay estimation; this research secured a $100K NMRC award to support model validation and translation.
- Supervised and guided NUS MSBA interns to deliver an optimization algorithm and discrete-event simulation model for ED waiting-time reduction; simulation results demonstrated a 37% reduction in mean waiting time.
- Engineered clinical natural language processing solutions using LLMs, BERT transformers, embeddings, and named-entity recognition, enabling advanced text mining of clinical documentation.
- Developed and optimized SQL pipelines from Epic Clarity to support predictive modeling, operational dashboards, and clinical research, across various departments within the institution.
- Earned multiple certifications: 5 Epic certifications, 2 AWS certifications (Cloud Practitioner, AI Practitioner), and Certified Kubernetes Application Developer (CKAD).

Land Transport Authority (LTA) Singapore **May 2021 - Jul 2022**

Assistant Manager, Ticketing Strategy and Future Projects

- Conducted rigorous quantitative analysis on transit ticketing and financial data to support evidence-based policy recommendations focused on improving financial sustainability and cost-effectiveness of transit ticketing systems and services. Secured a 5-year financial framework agreement between five stakeholders.
- Led as project manager and primary point of contact for an external business consultancy team during the execution of a comprehensive business review as part of Phase 1 study for merger of TransitLink and EZ-Link.
- Managed end-to-end ETL process using SQL to create dashboards in Tableau from transit ticketing and financial data, enabling division-level data-driven decisions and boosting operational efficiency.

Intern, Active Mobility Group **May 2019 - Aug 2019**

- Conducted a comprehensive study on the financial sustainability of Singapore's bike-sharing industry, leveraging stakeholder interviews and surveys to identify key policy levers, and built a System Dynamics simulation model to quantify policy impacts.
- Presented findings and recommendations before the Regulatory Bike-Share Implementation Committee, and subsequently on an international platform at the 2020, 3rd Asia Pacific System Dynamics Conference. Recognized as a paper of note by the System Dynamics Society official website.

EDUCATION

National University of Singapore **Aug 2022 - Jun 2023**

Master of Science in Statistics - (CAP): 4.06 of 5.00

National University of Singapore **Aug 2017 - Jun 2021**

Bachelor of Science with Honours (Distinction) - CAP: 4.21 of 5.00

Major in Applied Mathematics, Second Major in Economics

ACADEMIC PROJECTS

Optimization of LightGCN for Enhanced Recommender Systems **Jan 2023 - Apr 2023**
Capstone Project for M.Sc. in Statistics

- Devised novel strategies to mitigate over-smoothing in LightGCN Graph Convolution Network deep learning architecture, improving its performance in graph representation learning for recommender systems. Cluster-based sampling via community detection algorithms (Metis and Graclus) and dimensionality reduction was employed to achieve up to 14.2% improvement in model performance and significant reduction in training time.
- Awarded a top-three placement for project presentation within cohort.

Temporal Dependency Modelling via Time Series Imaging **Jan 2023 - Apr 2023**
Project for DSA5204 Deep Learning and Applications

- Investigated integration of time-series imaging using Gramian Angular Fields and Long and Short-Term Time Series Network (LSTNet) deep learning architecture for enhanced modelling of temporal dependencies in time series forecasting. Hyperparameter tuning yielded improvements in model test performance by approximately 30%.

PUBLICATIONS

Perioperative emergency laparotomy pathway for patients undergoing emergency laparotomy: A propensity score matched study **Dec 26, 2024**
The Annals, Academy of Medicine, Singapore

A prospective cohort study evaluated the EMergency Laparotomy Audit (EMLA) pathway in a single center, demonstrating significant reductions in surgical complications, length of stay, and costs, with improved protocol adherence but unchanged 30-day mortality. Findings support standardised perioperative pathways to enhance outcomes and inform EL care policies.

Development of an explainable machine learning model for predicting neurological deterioration in spontaneous intracerebral hemorrhage **Mar 25, 2025**
Intelligence-Based Medicine

Developed a machine learning model predicting early neurological deterioration in ICH, achieving an AUC-ROC of 0.8743. Combined multiple algorithms in a blended ensemble approach, outperforming individual models in prediction accuracy. SHAP analysis revealed key predictors (GCS, IVH, hematoma volume) consistent with clinical expertise. Utilizes readily available clinical and radiological data for real-time risk stratification in patients. Focused on early 48-h deterioration window, critical for timely intervention in Asian population.

VOLUNTEERING

National University of Singapore **Oct 2024 - Present**
Faculty of Science Postgraduate Alumni Ambassador

TECHNICAL SKILLS AND PROFESSIONAL CERTIFICATIONS

- **Languages/Tools:** Python, R, MATLAB, Git, PostgreSQL, Tableau, Bash Scripting, Docker, Kubernetes
- **Libraries/Frameworks:** TensorFlow, Scikit-Learn, PyTorch, Pandas, NumPy, Various packages for performing Time Series Analysis, Spatial Statistics, Network Analysis, Natural Language Processing, and building Generative AI applications (LangChain, Streamlit).
- **Machine Learning:** Supervised Learning (Regression, Classification, SVM, Decision Trees, Random Forest), Unsupervised Learning (Clustering, PCA), Gradient Boosting Methods (XGBoost, LightGBM)
- **Deep Learning:** DNNs, CNNs, Sequential Models (RNNs, LSTMs), GNNs
- **Big Data Tools and Cloud:** Apache Spark (SparkSQL, PySpark), Hadoop, AWS
- **Professional Certifications:** Text Processing (NUS, 2022), Data Structures and Algorithms (NUS, 2023), Machine Learning Operations (NUS, 2024), AWS Certified AI Practitioner (2024), AWS Certified Cloud Practitioner (2025), Certified Kubernetes Application Developer (CKAD) (2025)