KAUNG SI THU

Junior Data Scientist | Data Analyst

+60 178742670 | kaungsithu.sallius@gmail.com | linkedin.com/in/kaung-si-thu01

Junior Data Scientist with a strong foundation in data analysis, machine learning, and predictive modeling. Proficient in Python, SQL, Tableau, and Excel, with hands-on experience building end-to-end data science projects including feature engineering, model training, evaluation, and dashboard deployment. Passionate about transforming complex datasets into actionable insights and scalable data-driven solutions.

Work Experiences

Fireclash Enterprises - Rajkot, India   May 2023 - Dec 2023

Data-Driven Web Development Intern

- Utilized SQL to analyze user behavior and feature engagement data from a medical web application, enabling data-driven UI/UX enhancements.
- Designed and maintained a relational database schema for hospital patient records and bed management modules, reducing data retrieval time by 40% and improving real-time access to operational insights.
- Collaborated with a cross-functional team of developers and analysts, integrating hospital data pipelines into product workflows using Git and Agile, delivering 3 major feature updates.

Take Your Time - Yangon, Myanmar   Feb 2021 - Dec 2022

Technical Project Lead (Web & Data-Focused)

- Analyzed stakeholder requirements and conducted user research on over 100 feedback entries to define key data needs, content priorities, and information architecture for a nonprofit education website.
- Led a team of 3 using Agile and Git to deliver key features on time, resulting in a 30% boost in user engagement within the first month.

Education Level

Universiti Teknologi Malaysia – Johor Bahru, Malaysia   Oct 2025 - Expected Feb 2027

Master of Data Science

Marwadi University - Rajkot, India   Jun 2021 - Apr 2025

Bachelor of Technology in Information Technology

- Led a capstone project on unemployment forecasting, developing machine learning models in Python that achieved 90%+ predictive accuracy, supporting data-backed policy insights.
- Completed industry-recognized certifications (Google Data Analytics, DataCamp Data Scientist) while pursuing to strengthen skills in statistical analysis, machine learning, and data storytelling.

Technical Skills

- **Programming & ML:** Python (Pandas, NumPy, Scikit-learn), Machine Learning, MLFlow
- **Data & Databases:** SQL (PostgreSQL, MySQL), Data Modeling, Feature Engineering
- **Cloud & Big Data:** AWS, Apache Spark
- **Visualization & Tools:** Tableau, Excel (Advanced, PivotTables), Git, Docker

Projects

Unemployment Analysis and Prediction in India & Pakistan (*Team Leader*)   Mar 2025 - Apr 2025

- Analyzed unemployment records from India and Pakistan using Python (Pandas, Matplotlib) to identify macroeconomic trends and regional disparities.
- Developed regression and ensemble-based models to forecast unemployment rates with 90%+ accuracy, supporting data-informed policy recommendations.
- Created interactive Plotly dashboards to visualize insights and present findings to academic stakeholders.

Predicting HDB Resale Prices in Singapore   Jan 2025 - Feb 2025

- Cleaned and engineered features from 100,000+ rows of resale flat transaction data, incorporating town, flat type, lease, and geolocation (OneMap API).
- Engineered features from over 100,000 transaction records to build an XGBoost model that predicts HDB resale prices with over 90% accuracy, empowering Singaporean home buyers and sellers to make confident, data-driven financial decisions.
- Deployed results through a Plotly Dash tool that allowed real-time price prediction based on user-selected flat features.

Forecasting Tourism Demand in Singapore   Nov 2024 - Dec 2024

- Collected and integrated historical tourism arrival data, weather patterns, and event schedules to model demand fluctuations in Singapore's tourism sector.
- Developed a forecasting model using Random Forest and XGBoost (RMSE < 10%) to predict monthly tourist arrivals, delivering an interactive Plotly Dash dashboard that allows tourism stakeholders to proactively manage staffing, marketing spend, and resource allocation based on data-driven demand signals.

Additional Information

- **Certification:** https://drive.google.com/drive/folders/19Aq6dByi5FzzKfxc8rI0uPv3Aq8tN46v?usp=sharing
- **My Portfolio Website:** https://kaung-si-thu-the-analyst.vercel.app/