SU THET MIN HTEI













About Me

A freshly graduated Computer Science Student specializing in Big Data with a strong foundation in machine learning, data analysis, and programming. Completed hands-on projects, including building machine learning models and deploying web apps. Looking to leverage my knowledge and experience into a role in Data Science.

Education

University of Wollongong, Australia — Bachelor of Computer Science, Big Data

Mar 2023 - Mar 2025

- Current Weighted Average Mark (WAM): 73.94 / 100
- · Specialized in Big Data, focusing on data mining, machine learning, and scalable systems, Gained strong foundation in Python, SQL, data structures, and algorithm design through project-based learning, Developed real-world skills in data wrangling, visualization, and model evaluation using libraries like Pandas and Scikit-learn
- · Completed two capstone software development projects, demonstrating teamwork, communication, and technical delivery skills using agile methodology, Excelled in coursework involving distributed computing, data pipelines, and cloud-based data storage systems

Singapore Institute of Management — Diploma of Information Technology

Apr 2022 - Mar 2023

- · Gained foundational knowledge in computer science principles, including programming, databases, and systems development technique.
- · Learned to write code in programming languages such as Python and Java, the fundamentals of computer Networking, and IT Project Management.

Work Experience

Myanmar Apex Bank | Intern

July 2025 - Present

- Extracted and manipulated large datasets from Oracle databases using advanced SQL queries to ensure accuracy, relevance, and optimization for performance.
- · Collaborated with business and technical teams to design data logic, clean and transform raw data, and deliver automated reports supporting strategic decision-making.
- Enhanced financial data structure understanding by optimizing SQL queries and improving data pipeline efficiency.
- Developed and maintained dashboards to visualize key banking metrics, enabling real-time performance tracking.
- · Contributed to collaborative reporting processes in a fast-paced banking environment, improving cross-departmental data sharing.

Projects

Sentiment Analysis

Feb 2025 - Feb 2025

Personal Project

- Developed a sentiment classification model using XLM-RoBERTa (Hugging Face) to analyze tweets.
- Fine-tuned the model on social media data using tokenization and transfer learning, leveraging attention-based mechanisms to improve contextual understanding.
- Evaluated model performance using accuracy, F1-score, and confusion matrix, achieving over 90% accuracy.

Traffic Bottleneck Identification on Road Network

Oct 2024 - Feb 2025

Final Year Project

- Led a team of 4 as project leader to design and implement a dynamic traffic Bottleneck Identification system.
- · Applied Python and graph-based algorithms to detect congestion points and suggest alternate routes.
- Integrated map-based visualizations and real-time/synthetic traffic data analysis for bottleneck identification.
- Delivered geospatial analytics reports that supported traffic optimization and delay reduction strategies.
- Live Demo: FlowX App

CAERS Data Analysis

Aug 2024 - Aug 2024

Personal Project

- A Conducted exploratory data analysis on the CAERS dataset from Kaggle to uncover patterns in adverse food event reports.
- Cleaned, aggregated, and visualized data using Python (Pandas, Seaborn) to generate actionable insights.
- Built an interactive **Streamlit application** for dynamic exploration of trends in adverse event data.
- Live Demo: CAERSDataAnalysis Streamlit App

Skills

Languages: Python, Java, C++, SQL (NoSQL, PostgreSQL)

Data Visualization: Matplotlib, Seaborn

Libraries & Frameworks: Flask, Streamlit, Spark, Hadoop, Pandas, NumPy, Scikit-Learn, TensorFlow, PyTorch

Workflow Tools: Git, VS code, Jupyter Notebook, Google Colab

Languages

• English (Fluent) • Burmese(Native)