

AUNG KYAW WIN

DATA SCIENTIST & AUTOMATION SPECIALIST

Yangon, Myanmar | akw9785@gmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio Website](#)

EDUCATION

Bard College | Parami University (Dual Degree Program)

Aug 2022 - Jun 2026

Bachelor of Arts in Statistics and Data Science (SDS)

- CGPA: 3.55
- Honors/Awards: Magna Cum Laude, Rising Generation Summer Internship 2024 Recipient, Parami Undergraduate Scholar

PROFESSIONAL EXPERIENCE

The Water Agency

Oct 2025 - Jun 2026

Junior AI Specialist

Remote

- **Automated cross-functional workflows** using Python, n8n, and MS Power Automate, reducing manual processing time by 30% across core agency operations.
- **Architected and deployed an internal knowledge chatbot** leveraging Microsoft Copilot, **reducing organizational data retrieval times by 50-60%** and securing internal data access.
- **Spearheaded technical rollouts and a monthly AI workshop series** for non-technical employees, directly increasing internal AI tool adoption and cross-departmental productivity.
- **Built modular engineering pipelines** for large-scale grant discovery, knowledge retrieval, and research synthesis, **cutting time-to-insight by 50%** and accelerating project proposal pipelines

Synergy Solutions

Jul 2025 - Sept 2025

Data Analyst Intern

Remote

- Streamlined educational processes for nearly **200 students through a centralized digital classroom workflow**.
- Integrated **Google Sheets and Google Drive APIs** for real-time assignment management.
- Implemented an **automatic flagging system** for overdue submissions.
- Designed **Google Apps Script automations** to manage workflows seamlessly.
- Created data flow diagrams to facilitate educators' adoption of the new system.
- Developed **an automated pipeline to transfer Zoom recordings to Google Drive**.
- Ensured all meeting data was automatically updated in a centralized Google Sheets dashboard.
- Significantly reduced manual tasks, enhancing efficiency and data visibility.

The Water Agency

Jun 2025 - Aug 2025

Data Analyst Associate

Remote

- Took ownership of a disorganized and manual LMS and website data system by **building a Python Selenium automation tool to extract raw data from the website admin panel**
- Designed a modular **data pipeline in Python to clean, transform, and generate standardized CSVs** from inconsistent raw files, enabling seamless integration with Power BI
- Built a multi-page Power BI dashboard featuring **20+ KPIs, slicers, models, and advanced visuals, enabling dynamic filtering across time, demographics, and course performance indicators**
- Delivered a **fully automated reporting system** where any team member can **generate an updated dashboard by simply rerunning the Python script** and refreshing the Power BI file, eliminating repetitive manual work
- **Reduced reporting turnaround time by over 80%** by streamlining manual data collection, cleaning, and dashboard refresh processes into a scalable, script-based system

Cassidy Levy Kent LLP

Jun 2024 - Aug 2024

Data Science and International Trade Research Intern

Washington D.C., USA

- Conducted in-depth research on various HTS Codes and Tax laws
- Authored a comprehensive case study report on **Section 301 tariffs, focusing on semiconductors and EVs, providing detailed insights into the economic and legal impacts**
- Contributed code to the development of programs used by the firm's data scientists to analyze production and financial data relevant to trade remedy cases, improving efficiency and accuracy
- **Automated data extraction from PDF Files using Python**, converting and formatting extracted data into accessible formats, streamlining data workflows, reducing manual processing time, and improving productivity

Parami University ~ Google Career Certificates Program, Remote

Dec 2023 - Jun 2024

Program Assistant

Remote

- Supported admissions processes by **evaluating applications and managing communications**, ensuring a seamless onboarding experience
- Assisted and helped **facilitate online interviews and information-sharing sessions for over 200 students, hosting 20-30 students** in each session
- Facilitated **monthly online learning sessions and tracked student progress** on Coursera, achieving a higher completion rate for enrolled participants

ASEAN Foundation, Remote

Mar 2023 - Jun 2023

ASEAN Data Science Explorers Ambassador

Remote

- Successfully collaborated with multiple universities, expanding the reach of the initiative
- Designed and published promotional materials to encourage student registrations, ensuring maximum reach and engagement
- Delivered impactful workshops to **around 300 students (double the requirement of the program)**
- **Facilitated interactive sessions with 50 - 100 students** per session for in-person sessions, delivering data science concepts
- Oversaw and gave feedback to student projects, resulting in **7/10 teams selected for National Finals being taught by me**

PROJECTS

Candlestick Pattern Image Classifier - Python, PyTorch, Vision Transformers, Gradio

[GitHub](#) | [Demo](#) | [Medium Article](#)

- Developed a **deep learning model to classify key candlestick patterns** (Doji, Hammer, Engulfing) from stock chart images, automating technical analysis traditionally done by traders
- Collected and processed historical stock data using yfinance and pandas; generated thousands of labeled candlestick chart images with mplfinance and custom cropping scripts
- Designed and trained a custom **Vision Transformer architecture on a curated dataset**, achieving robust classification accuracy without data augmentation
- **Deployed an interactive web app with Gradio** on Hugging Face Spaces, enabling users to upload candlestick images and receive real-time pattern predictions
- Learnt valuable lessons on data quality, model simplicity, and deployment, making the project publicly accessible to both traders and ML scientists

Financial Tweet Sentiment Classifier – Python, Hugging Face Transformers, PyTorch, Gradio

[GitHub](#) | [Demo](#) | [Medium Article](#)

- Built a state-of-the-art **NLP model fine-tuned on a specialized financial tweets** dataset to classify sentiment as bullish, bearish, or neutral
- Leveraged Hugging Face's cardiffnlp/twitter-roberta-base-sentiment-latest transformer for tokenization and fine-tuning using the Trainer API, achieving high accuracy on unseen tweets
- **Automated data preprocessing, including tokenization and dataset splitting**, with the Hugging Face Datasets library
- Developed and **deployed an interactive Gradio web app for real-time sentiment prediction**, making the tool accessible to traders and researchers
- Published the trained model and demo on Hugging Face Hub, enabling easy sharing and integration

Instagram Reels Recommendation System – Python, Graph Theory, Hash Maps, Streamlit

[GitHub](#) | [Demo](#) | [Medium Article](#)

- Designed and implemented a **multi-strategy personalized recommendation engine combining graph-based social filtering, content-based filtering, collaborative filtering, and popularity boosting for short-form video content**
- Generated a **synthetic dataset simulating 100 users and 300 reels** with realistic engagement metrics to test and refine algorithms safely without real user data
- Employed **efficient data structures (hash maps, adjacency lists) for fast O(1) lookups** and optimized graph traversal to scale social connection analysis
- Developed detailed, explainable recommendation outputs highlighting social connections, tag preferences, and content popularity to improve user trust and engagement
- Built an interactive Streamlit web app and command-line demo for exploring personalized recommendations with tag distributions and social network visualizations
- Focused on system performance through caching, weighted scoring, and modular design; designed for future improvements such as time decay and real-time updates

VOLUNTEER/EXTRA CURRICULAR ACTIVITIES

- Video & Photo Editor of student-led Voice Acting Club ~ Parami University
- Communications Officer of student-led Datalliance Club ~ Parami University
- Video & Photo Editor of student-led Media Club ~ Parami University
- Web Designer of student-led LAPC Club ~ Parami University
- Voice Donor ~ Myanmar Christian Fellowship For The Blind

CONTINUING EDUCATION COURSES

Technical Support Fundamentals, Data Analytics ~ Google (2022)

Introduction To Computer Science ~ Harvard University (2022)

Learn To Program: The Fundamentals ~ University of Toronoto (2023)

AI Engineer for Data Scientist Associate Certification (In Progress) ~ DataCamp (In Progress)

TECH STACK

- **Programming Languages:** Python, R, SQL
- **Data Analytics Tools:** PowerBI, SAP Analytics Cloud, Tableau, QGIS

AWARDS/ACTIVITIES

- ASEAN Data Science Explorers 2022 (2nd Runner-Up)
- Rising Generation Summer Internship Program 2024 Recipient
- State-level Mathematics Champion 2017
- Participant in the International Essay Contest for Young People (2020)

ADDITIONAL INFORMATION

Languages: English (Proficient), Burmese (Native), Hindi (Native), Bihari (Native)