

VALENTINO ONG

<https://valent2809.github.io/Valentino.io/> | valentinoong61@gmail.com | +6591642003 | [linkedin.com/in/valentino-ong/](https://www.linkedin.com/in/valentino-ong/)

WORK EXPERIENCE**AI Intern, NCS****May 2024 – Aug 2024**

- Led a project from an overseas client to automate tender analysis process using Gen AI from scratch
- Achieved a 95% reduction in time taken for manual processing of 1000+ pages of documents
- Projected cost savings of over \$50,000 by reducing manual effort and increasing operational efficiency
- Collaborated with stakeholders for requirements gathering and presented project plan and performance insights across use cases, leading to client approval.
- Designed and developed the full-stack solution, leveraging Retrieval-Augmented Generation (RAG) framework and advanced prompt engineering techniques to deliver accurate and efficient outcomes

Data Scientist Intern, Galen Growth**Jun 2023 – Aug 2023**

- Achieved 90% accuracy in classifying news and companies from unstructured data using machine learning models (multi-label classification), eliminating manual work entirely.
- Develop automated data web scraping scripts in Python using BeautifulSoup and Selenium, achieving an 80% reduction in time spent extracting data from online sources.
- Statistical analysis, processed, cleaned, and formatted various data using Numpy, Pandas, MySQL and Excel
- Designed database schema and implemented ETL pipeline to insert thousands of data into database

EDUCATION**SINGAPORE MANAGEMENT UNIVERSITY****Aug 2021 – Apr 2025****Bachelor of Computer Science (AI & Business Analytics)**

- **Honors:** Magna Cum Laude / High Distinction
- **GPA:** 3.69 / 4.00
- Dean's list AY24-25 & AY23-24, Trailblazer ASEAN Scholarship -2024, 2023, 2022 and 2021
- Teaching Assistant (Intro to AI & Analytics Foundation)

PROJECT**AIP National AI Student Challenge 2024 ([GitHub Repository](#))**

Shorted listed for Final stage in track 1: AIP Association challenge problem for pet adoption rates prediction

- Performed EDA on PetFinder.my data to clean and filter dataset of 14,993 rows of data and 49 features
- Built an end-to-end Machine Learning pipeline that include data preprocessing and utilizes machine learning algorithms such as XGBoost and Random Forest to predict pet adoption rates

Smart Energy Analysis using AWS & GenAI

Scalable ETL and analytics pipeline to optimize solar energy adoption strategies using big data

- Built an automated big data ETL pipeline using AWS Glue and PySpark to process millions of records
- Applied clustering using AWS SageMaker to segment customer groups to identify high-potential targets for solar energy adoption strategy optimization
- Enabled retail strategy optimization by delivering insights through interactive GenAI chatbot for stakeholder

Social Analytics of reviews on Lazada ([GitHub Repository](#))

Scrape & analyse Lazada reviews from App store, Google play & other sites to provide actionable insights

- Identified insights on customer satisfaction, pain points, and potential strategies for enhancing app retention
- Performed EDA: data cleaning, timeline analysis, and word clouds to identify key trends in reviews
- Conducted sentiment analysis with VADER to assess customer opinions and validate rating accuracy
- Applied Topic Modelling (LDA) to uncover key themes and areas of improvements in reviews

LinkedOut: Resume and Job Matching Recommendation System

Built a recommendation system using Doc2Vec and FAISS to match resumes with job descriptions and vice versa

- Developed a semantic search model for candidates to find suitable jobs and recruiters to find ideal candidates
- Leveraged FAISS for similarity matching and Doc2Vec embeddings to capture contextual meaning from text
- Developed the backend REST APIs using FastAPI for resume-to-job and job-to-resume recommendations

SKILLS & INTERESTS

- **Certification:** Udemy Data Science and Machine Learning bootcamp, Udemy Taming Big Data with Apache Spark and Python, Google Career Data Analytics, AWS Certified Solutions Architect – Associate
- **Programming Languages:** Python, SQL, Java & C
- **Data Science & Analytics:** Data Mining, Machine Learning, Gen AI, PySpark, NLP, Web scraping & Tableau
- **Web Development/Frameworks:** FastAPI, Streamlit & Spring Boot
- **DevOps & Deployment:** Docker, Kubernetes, CI/CD pipeline development, AWS
- **Language fluency:** Effectively bilingual in English and Indonesia, Chinese (A-level) & Malay