Michael Geraldin Wijaya

Jakarta, Indonesia | +62-895-6043-39001 | michaelwjy2@gmail.com
linkedin.com/in/michaelgeraldinwijaya | github.com/Michs224 | michgerald.tech

EXPERIENCE

ID/X Partners, Jakarta, Indonesia

Project-Based Virtual Intern: Data Scientist – ID/X Partners x Rakamin Academy 08/2023 – 09/2023

- Finishing various task related with the activity of Data Scientist from ID/X Partners, such as Big Data Fundamental, Statistics & Data Analytics, and Machine Learning Implementation.

EDUCATION

Bina Nusantara University
Bachelor of Computer Science
Streaming: Intelligent System
Computer Science, 2022—Present Latest GPA: 3.94
Awards:
- Awardee Widia Partial Scholarship BINUS University
- Awardee of Beasiswa Unggulan Kemendikbudristek 2023
- National Finalist ASEAN Data Science Explorers 2024 Competition
- Top 20 Finalist School of Computer Science Hackaton 2024 BINUS x Microsoft (AI4A)

Senior High School 4 Palembang
Natural Science, 2019—2022 Grade: 97.29
Awards:
- 1st Place Winner of the National Independent Science Competition - Mathematics.
- A+ (Gold Medalist) in the Indonesian Science Competition 2022 - Mathematics.
- A+ (Gold Medalist) in the Indonesian Science Competition 2022 - Technology and Information.
- Gold Medalist in the 2022 National Numeracy Olympiad for High School Students.
- Gold Medalist in the National Student Science Competition - Mathematics.

PROJECTS

Patient's Condition Classification Using Drug Reviews 06/2024

- The project uses drug reviews with machine learning models such as Passive Aggressive Classifier, Random Forest, and XGBoost to predict patient conditions. It applies text processing techniques such as CountVectorizer, TFIDFVectorizer, and Word2Vec to analyze user reviews. The project provides visualizations of condition proportions and word clouds, and recommends the top three drugs for the condition.

Personal Medical Chatbot with LLM RAG 06/2024

- Personal medical chatbot project using Large Language Models (LLM) and Retrieval Augmented Generation (RAG). The chatbot provides detailed answers by leveraging medical knowledge from PDF documents, employing LlamaCPP for LLM integration and llama_index for efficient document handling. This approach enhances natural language responses and enables advanced retrieval of health-related information.

Simple Personal Chatbot 06/2024

- This simple personal chatbot project uses NLP techniques and neural networks to process user queries effectively. It employs TF-IDF and Word2Vec to enhance comprehension and responsiveness. Data preprocessing techniques like tokenization, punctuation removal, and stop word elimination help the chatbot learn from an intents.json dataset to categorize patterns and responses. TF-IDF evaluates word importance, while Word2Vec creates word vectors for better semantic understanding. Flask is used to develop API endpoints, enabling integration with various platforms.

Ad Hoc Insights Consumer Goods 06/2024

- This project aimed to analyze consumer goods data to provide insights for management decisions. Using MySQL, I implemented concepts such as Joins, Subqueries, String, Numeric, and Date Functions, as well as Window Functions and Common Table Expressions (CTEs) for effective data manipulation. To visualize the findings, I used Tableau to create clear and informative data visualizations, and Canva for presentation creation.

LibraLink: Smart Attendance with Book Recommendation 09/2023 – 12/2023

- LibraLink is an advanced attendance system integrated with a book recommendation feature, specifically designed for library environments. At its core lies a robust facial recognition-based attendance mechanism, powered by Convolutional Neural Networks (CNN). In addition to tracking attendance, the system offers a personalized book recommendation service, leveraging the K-Nearest Neighbous (KNN) algorithm to align with user preferences.

Movie Recommender System 08/2023

- This ML project utilizes a Content-Based Recommender System for movies, using features like titles, genres, and posters to provide recommendations based on the chosen film. The system employs cosine similarity to gauge similarities between movies and is deployed using the Streamlit platform, enabling user-friendly interaction.

Image Classification of Billionaires 08/2023

- Utilized OpenCV for face and eyes detection, performed data cleaning through OpenCV's face detection, and conducted feature engineering using wavelet transforms. Built models using SVM, logistic regression, and random forest, further refining them with GridSearchCV. Exported final model to a file and established a Python Flask server. Deployed the project on AWS EC2.

Credit Risk Modelling 08/2023

- Utilized Logistic Regression to identify loan default risks from historical loan data. Conducted data preprocessing, feature engineering, and optimized model performance via ROC curve analysis. Gained insights into credit default risk factors, including loan duration, employment length, and payment history.

CERTIFICATES

- *Data Science and Machine Learning Fundamentals*
- *Machine Learning with Python*
- *Statistics Fundamentals*
- *Data Analysis using Python*
- *Python for Data Science*
- *Supervised Machine Learning: Regression and Classification*
- *ASEAN Data Science Explorers 2024 National Final Indonesia*

SKILLS

- Programming: Python, SQL, C, Java
- Data Visualization
- Machine Learning: Machine Learning Algorithms and Techniques
- Deep Learning: Tensorflow, Keras
- Natural Language Processing (NLP)
- MLflow
- Git/GitHub

Languages:

- Indonesian (Native)
- English (Intermediate)

ORGANIZATIONS AND ACTIVITIES

- Data Science Club (DSC) and Ureeka BINUS University (2024 - Present)
- Badminton (2023 - Present) and Binusian Gaming (2022 – Present) BINUS University
- Tutor BINUS University (2023 – 2024)

VOLUNTEERING

- Anti Bullying Educator - Christian Eunike Orphanage Foundation, Semarang, Central Java (11/2022)
- Patriotism Educator - Bethlehem Salvation Army Orphanage, Semarang, Central Java (06/2023)