

Muhammad Fahmi Bin Mohd Zainal

Kuala Lumpur , Malaysia
fahmizainal9@gmail.com • +60(165507849)
[LinkedIn Profile](#) | [Portfolio](#)

Summary

Data Scientist with 1+ year of experience in IT, data analytics, and data science, proficient in Python, SQL, Tableau, ETL, EDA, API Testing, Web Application , machine learning ,deployment and AWS Devops Fundamental. Successfully led projects with proven track record improving data processing efficiency by 90% and revenue growth by 10% through Survey Automation, predictive modeling and data-driven insights. Skilled in developing interactive dashboards, designing ETL pipelines, creating API Endpoints and maintaining and deploying to data warehouses such as Azure and AWS.

Professional Experience

INVOKE SOLUTIONS

Data Scientist

Kuala Lumpur
November 2023 – Now

Surveys Making, Automation and Optimization:

- Led the design and implementation of the [IVR Data Cleaner, Questionnaire Definer, and Keypress Decoder](#), significantly improving the efficiency of data processes from 8 hours to 30 minutes work.
- Spearheaded and handled survey projects for call sampling, [weighting](#) and [visualization through crosstabs and charts](#) utilizing google colab, Azure Databricks and in-house web applications.
- Developed and integrated [Unified Survey web applications](#) using Streamlit, enhancing collaboration by merging functionalities with team-developed applications in Streamlit and utilizing FastAPI, supported by Docker,AWS Services and GitHub Actions (CI/CD).
- Architected data for [Inisiatif Pendapatan Rakyat \(IPR\)](#) utilizing data mining techniques and Geocoding API to map the locality (latitude and Longitude) and postcode.

Machine Learning Development:

- Prepared Data ,Pipelines,Designed and developed predictive models to analyze [ROAS Benchmark and Campaign Benchmark,consumer behaviors, and market trends](#) utilizing MongoDB,Databricks,Azure Blob Storage for workflow and using advanced matching and machine learning techniques thus enhancing company's revenue up to 10% by engaging potential regular clients.
- Created innovative EDA and model deployment to predict [Personal Attributes such as Salary and Income Group](#) ,improving demographic targeting accuracy up to 90% accuracy.

ROAS Dashboard Application:

- Streamlined and innovated solutions for the [ROAS Dashboard application](#) to deliver more precise and filtered results. Enhanced interactivity with users by providing clear graphs for visualization.
- Worked with Frontend Developer and interacted by creating API Endpoints of the functions to pass the data as of difference in packages usage.

Computer Vision Applications Maintenance:

- Maintained [computer vision's project of FGV oil palm company](#) to optimize the pesticide application.
- Maintained [document parsing automation](#) by implementing TensorFlow and YOLOv5 technologies, facilitating the digitization of invoices and receipts into structured formats for detailed analysis.

Data Warehouse Management:

- Oversaw the maintenance of the INVOKE's data warehouse, ensuring robust integration and customization of dashboards to optimize data analytics capabilities.

EXCELERATE ASIA

Data Analyst Apprentice

Kuala Lumpur
March 2023 – October 2023

- **Workforce Dynamics Insights Project:** Enhanced data integrity and workforce management through expert data wrangling with Excel and development of Interactive Tableau Dashboards, providing crucial insights into [Retention, Compensation Fairness, and Diversity](#).
- **Global Superstore SQL Project:** Leveraged PostgreSQL for robust data management and extraction to Tableau, analyzing trends in product returns to support strategic business decisions thus completing analytics bootcamp training.
- **ML Project** Achieved top 28% ranking out of 1908 teams in a [Binary Prediction Competition On Smoker Status Using Bio-Signals](#), leveraging logistic regression, random forest, and tree-boosted algorithms such as Gradient Boosting, LightGBM, XGBoost, and CatBoost..
- Planned the work meticulously, conducting meetings to discuss strategies, optimize models using Optuna for hyperparameter tuning, and implement a weighted voting classifier for enhanced performance, resulting in an ROC-AUC score of 0.87178 which is 87% accuracy.

- **Radiation Studies:** Conducted thermoluminescence research on microscope glass slides using Fourier transform infrared (FTIR), Raman, and Photoluminescence (PL) spectroscopies after exposure to 60Co gamma-rays. Enhanced understanding of radiation effects, improving dosimeter applications in safety measurements.
- **Reactor Maintenance and Safety:** Played an integral role in the maintenance of the TRIGA PUSPATI reactor, contributing to safety tests and the neutron flux project. Gained significant project management skills and deepened nuclear industry insights through hands-on experience.

Education

UNIVERSITY OF MALAYA
Bachelor of Science in Physics (Hons)

Kuala Lumpur
September 2019 – February 2023

- University Book Prize Award recipient
- 3 times Dean's List Academic Achievement Award recipient
- First-class degree Honors

MALAYSIAN ARMED FORCE
Second Lieutenant Reserved Army

Kuala Lumpur
September 2019 – February 2023

- **Leadership:** [Commanded a platoon of 89 cadets](#), achieving high training ratings and improving administrative efficiency, reducing errors by 5%.
- **Communication:** Excelled in high-pressure scenarios, coordinating team efforts and earning recognition for leadership during critical simulations.

Certifications

Advanced LLM Certificate, 2024
Oil & Gas Meta HRDC Fundamentals Of Oil and Gas Software for Professional – Power BI, Advanced Excel, Primavera P6, Microsoft Project, AutoCAD , SAP Fundamental, 2024
General Assembly Data Analytics and Data Science Bootcamp, 2023
General Assembly Data Analytics and Data Science Bootcamp, 2023
Kaggle Python certification, 2023
Coursera Working with Big Query certification, 2023
IBM Data Analytics Essential Certificate recipient, 2023
IBM Excel Essential for Data Analysis Certificate recipient, 2023
IBM Data Visualization and Dashboards with Excel and Cognos Analytics Certificate recipient, 2023

Skills

Software and Technologies (Professional Working Proficiency)

- Programming Languages: Python, Pyspark ,SQL, R ,HTML5,CSS3, PHP Fundamentals, JavaScript Fundamentals
- Cloud & Databases: Azure Databricks,Azure Functions, AWS Services (ECR, EC2, API Gateway, Fargate, S3, Lambda ,EventBridge ,Sagemaker), GCP Fundamental, MongoDB
- Data Processing and Analytics: PySpark, Google Colab, Streamlit Web Application, OriginPro
- Data Science Packages : Pandas, Numpy, BS4, Sklearn, TensorFlow, Keras, Matplotlib, Seaborn
- Data Science Algorithms: Linear Regression, CatBoost ,LightGBM, XGBoost, Random Forest , Logistic Regression, KNN, SVM , CNN and K-Means Clustering ,Binary Classification ,Linear Regression.
- Dashboarding and Reporting: Microsoft Office (Excel, PowerPoint, Word), PowerBI, Tableau, Google Sheets, Google Sheet Dashboard, Google API Sheets, Microsoft Excel VBA Macros
- Web Development: FastAPI, Streamlit Web App
- Containerization and Virtualization: WSL , Ubuntu ,Docker
- Development Tools: Visual Studio Code, GitHub Actions (CI/CD)