

# Sulaiha Subi

## Data Scientist

ssulaihasubi@gmail.com | (+60) 10-9418598 | [linkedin.com/in/sulaihasubi](https://www.linkedin.com/in/sulaihasubi) | [github.com/sulaihasubi](https://github.com/sulaihasubi) | [sulaihasubi.github.io](https://sulaihasubi.github.io)

### PROFESSIONAL EXPERIENCES

---

#### MARQETR

Kuala Lumpur, MY

#### Data Scientist

Jan 2022 – Present

- Led multi-disciplinary 7 person team to design, develop, and launch MARQETR 2-sided on demand tech enabled platform that operating like the “uber of marketing”, powered by a smart-matching algorithm
- Extend the knowledge and skills in Data Science for marketing incorporating Machine Learning & Smart Matching Algorithm
- Developed strategic insights for MARQETR, including marketing strategy, end-to-end data pipeline and operations departments
- Design training and demo for AirAsia Academy Digital Marketing student engagement targeted at Fresh Grads and upskill candidates; led the training & demo to be on boarded to the platform that were attended by over 200 students within 3 months
- **Achievement:** Successfully deploy MARQETR Smart Matching Algorithm by help businesses save time and money, and be in control of their marketing cost and resources by providing access to vetted and verified on-demand high-skilled marketers

#### Project Experience

- MARQETR Smart Matching Algorithm
  - Developed Smart Matching Algorithm in use to matches more than 160 marketers with 20+ business users resulting in reduction of 20% in save time and money for businesses by providing access to vetted & verified on-demand high-skilled marketers
  - Developed end-to-end data science workflow by identifying & define the problems, building data pipelines for MARQETR, deploy & validate the solution
  - Generate actionable insights for business improvements
  - Managed and assess the effectiveness and accuracy of new data sources and data gathering techniques to improve and consolidate end-to-end processes the way businesses connect and collaborate with marketers

#### RISEHILL DATA ANALYSIS

Kuala Lumpur, MY

#### Data Scientist

Jun 2021 – Dec 2021

- Ensure the data quality by performing data cleaning, normalization & enrichment
- Working on every aspect of Data Analytics and Predictive Modeling
- Develop data models and algorithms to apply to data sets
- Used Machine Learning algorithms & statistical techniques to produce solutions to the business problems
- Conduct research from which to develop prototypes and proof of concepts (POC)
- Built ML workflow & designated experiments to supply tailored reports to the manager, and potential clients
- Applied cutting-edge Computer Vision skills for POC in face recognition & counting for productivity measurements
- Built and test Machine Learning model in Dataiku DSS Platform to select the most appropriate ones for use on a project
- Develop new client relationships/potential collaboration
- Meet & present the company's product/services & highly responsible to prepare high business value commercial proposals across different domains of the industries (client-facing skills)
- **Achievement:** Utilize extensive knowledge in Data Science & Computer Vision such as developed predictive modeling, abilities to extract, clean, and prepare data for data science model ingestion and communicate those complex & technical outputs to the business peoples or industry experts

#### Project & Publication Experiences

- [Integrated Data Intelligence \(IDI\)](#)
  - Developed & maintained NLP features in IDI: Cleaning & preprocessing text for oil and gas documents such as pdf, PowerPoint, .las, etc. with SpaCy
  - Resulting less time in searching by advanced searching with fast 3D browsing from millions of files

- **Trained Machine Learning Models (Auto-ML) with Dataiku DSS (POC)**
  - Preprocessed 4993 online transaction records, performed feature engineering from past customer online transaction
  - Trained Customer Transaction Fraud Detection Model, achieving 0.958 ROC AUC score
  - Built Insurance Claim Fraud Detection from Kaggle Dataset
  - Present those ML models to the industry experts, clearly communicate complex & technical information
- **Trained Machine Learning Models (Auto-ML) with Dataiku DSS (POC)**
  - Preprocessed 4993 online transaction records, performed feature engineering from past customer online transaction
  - Trained Customer Transaction Fraud Detection Model, achieving 0.958 ROC AUC score
  - Built Insurance Claim Fraud Detection from Kaggle Dataset
  - Present those ML models to the industry experts, clearly communicate complex & technical information
- **Face Recognition and Counting for Productivity Measurement for Drone Company (POC)**
  - Built face recognition and count the number of people based on the number of individuals detected entering the designated region
  - Calculate dwell time (duration) of each detected person for time spent in the designated region
  - Built Face Recognition using OpenCV with overlay bounding box on each visible individual smoothly & continuously
- **Dataiku Frontrunner Awards (2021)**
  - Use Case Submitted: Using AI to combat the Rise in Corporate Fraud Malaysia
  - The only use case submitted from APAC regions which is represented Malaysia
  - More details about the use case can be read here:  
<https://community.dataiku.com/t5/Dataiku-Frontrunner-Awards/RiseHill-Data-Analysis-Using-AI-to-combat-the-Rise-in-Corporate/ta-p/18184>

## OMDNA

### Machine Learning Engineer (Contributor)

New York, NY

Sep 2021 – Dec 2021

- One of the collaborators out of 47 AI Engineers from 18 Countries (8 weeks challenge)
- Build an automated plant health prediction model, using drone-derived multispectral and thermal data combined with weather and irrigation data
- Train a model to detect water stress, pest infestation, and man-made plant stress

### Project Experience

- **Automated Plant Health Prediction Model**
  - Build an automated plant health prediction model, using drone-derived multispectral and thermal data combined with weather and irrigation data (*github repository for this project is private due to TnC of the project*)

## PULSY TECHNOLOGY

### Associate Software Engineer

Cyberjaya, MY

Apr 2021

- Monitor and maintain the current system of Novartis and CBRE such as debugging and keep eye on the functionality of the systems
- Monitor AWS Management Console of Novartis and CBRE to manage their resources and well functions
- Provide documentation for the design, development, integrations, and deployment of the web app
- Work with IoT Engineer to perform and design testing protocols/prototypes to ensure that the product is fully tested before the deployment
- Ensure that all enhancements, bug fixes, and tasks assigned are well documented
- Able to teach and presents the programming or back-end part to other team members to ensure they understand how the code works and functions at the back-end side
- Contributor to the team member to ensure the project requirements are met

**Terra Weather (Apprenticeship)**  
**Deep Learning Infrastructure Development**

**Singapore, SG**  
Sep 2020 – Jan 2021

- Part-time Apprenticeship
- Develop new TW Caffe layers enhance data processing pipeline
- Optimize the TW Caffe backend
- Develop new layers using C++ and Smojo
- Extend the Willow of Smojo-to-C compiler
- Integrating open-source framework, ZeroMQ to streaming real-time I/O using C++

**Shopee**  
**Digital Marketer**

**Cyberjaya, MY**  
Sep 2020 – Apr 2021

- Product Research & Product Sourcing
- Manage E-Commerce Store & Run Facebook campaigns for each product
- Analyze marketing data and reporting
- Manage supply chain

**SkyMind**  
**Junior Deep Learning Engineer (Trainee)**

**Kuala Lumpur, MY**  
Apr 2020 – Jun 2020

- Trained by SkyMind's experienced Deep Learning Engineer on how to effectively use Deep Learning tools to create Computer Vision Models for productions
- Understand the state of the art of advanced machine/deep learning concepts, theory, and applications Deep Learning model development from deploying deep learning to production
- Hands-on with a large data set, learn algorithms, build models, and verify them
- Improved (fine-tune) deep learning models
- Knowledgeable in how to image processing, image classification, object detection to image segmentation, using the Deeplearning4j framework
- Knowledge in machine/deep learning, optimization, probability theory, and statistical inference
- Powerfully used programming in Java using IntelliJ IDEA and Python
- **Achievement:** Utilize extensive knowledge in deep learning to complete a Computer Vision Project (Face Recognition) within two days

**Project Experience**

- **Facial Recognition: [Are you in CERTIFAI Online CDLE?](#)**
  - Build a computer vision model using IntelliJ IDEA to detect and recognised faces of individual in real time of members in CERTIFAI Online CDLE where the members clustered into two group which consist of the trainers and the trainee
  - Resulting less time in collecting the trainers and trainee attendance
  - Tweak and enhanced the performance of the face recognition in the current model

**Department of Information System**  
**Developer (Internship)**

**Kuching, SWK**  
Feb 2017 – July 2017

- Debugged Service-Learning System for Local Support
- Performs upgrades and maintenance to the Service-Learning System
- Information and data gathering to be inserted into the Service-Learning System
- Managed the database of SL System
- Develop a website for the promotional and latest information about SL Projects including contents writing & web designing
- Performed research about SL & communicated the research both internally & externally to the Service-Learning's Advisor
- Engaged in team collaborations to meet the goals for implementation of Service Learning
- **Achievement:** Utilize extensive knowledge in a programming language such as PHP, HTML, MySQL, MySQLi, CSS, & database management to complete the SL System & SL website within the time frame

## EDUCATION

---

**UNIVERSITY OF MALAYSIA SARAWAK**  
*Bachelor of Computer Science, Major in Information System*

**Sarawak, MY**  
**2014-2019**

## CERTIFICATIONS

---

**DEEP LEARNING IN COMPUTER VISION**  
CERTIFAI

**SkyMind**  
30 April 2020 – 30 April 2022

**TAEKWONDO BLACK BELT 1<sup>st</sup> DAN**  
UNIMAS Taekwondo Club

**Taekwondo Malaysia (TM)**  
2018

## SKILLS

---

**Programming:** Python, JSON, HTML, HTML5, CSS, PHP, MySQL, NoSQL, Java, C++, R, Java Script, Prolog, XML, LaTeX

**NLP:** Named Entity Recognition (NER), Text Pre-Processing using SpaCy

**Computer Vision:** Face Recognition, Video Analysis, Image Processing, Optical Character Recognition (OCR)

**Tools:** IntelliJ IDEA, Adobe Dreamweaver CC 2017, Microsoft Visual Studio, Tableau, Weka, Eclipse, Notepad++, Rapid Miner, TeXstudio, CUDA, Maven, Laravel, Firebase, Dataiku, PyCharm, Jupyter Notebook & VS Code

**Others:** AWS EC2, Dataiku, Tableau, SAS, DAGsHub, Data Version Control (DVC)

## REFERENCES

---

### **Mawarni Tan Sri Adam**

Founder & CEO, MARQETR

*(contact number available upon request)*

[mawarni@marqetr.com](mailto:mawarni@marqetr.com)

### **Muhammad Zul Fadhli Bin Mohd Fauzi**

Geoscientist, RiseHill Energy Solution

(+60)11 - 219 25248

[muhammad.zul@rdagroups.com](mailto:muhammad.zul@rdagroups.com)

*further references available upon request*