

# Sulaiha Subi

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

## EXPERIENCES

---

### RiseHill Data Analysis

Kuala Lumpur, MY

#### Data Scientist

Jun 2021 – Present

- Ensure the data quality by performing data cleaning, normalization & enrichment
- Working on every aspect of Data Analytics and Predictive Modelling
- 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)

### Omdena

New York, NY

#### Machine Learning Engineer (Part-Time)

Sep 2021 – Present

- 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
- Tools: QGIS, DAGsHub, GoogleColab, DVC

### PULSY Technology

Cyberjaya, MY

#### Associate Software Engineer

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)

Singapore, SG

#### Deep Learning Infrastructure Development

Sep 2020 – Apr 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 data and report
- 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
- Used advanced techniques and algorithms of Deep Learning for Computer Vision
- 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 (meet the project timelines)

## PROJECTS & COMPETITIONS

---

### Dataiku Frontrunner Awards (2021)

- Use Case Submitted: Using AI to combat the Rise in Corporate Fraud 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>

### Integrated Data Intelligence (IDI)

- Create & maintain: Cleaning & preprocessing text for oil and gas documents such as pdf, PowerPoint, .las, etc. with SpaCy

### Trained Machine Learning Models (Auto-ML) with Dataiku DSS

- 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

### Face Recognition and Counting for Productivity Measurement (POC)

- Built face recognition and count the number of people based on the number of individuals detected entering the designated region
- Built Face Recognition using OpenCV with overlay bounding box on each visible individual smoothly & continuously

## EDUCATION

---

### UNIVERSITY OF MALAYSIA SARAWAK

Bachelor of Computer Science with Honours (Information Systems)

Kuching, SWK

2019

### LABUAN MATRICULATION COLLEGE

Biology, Chemistry, Computer Science

Labuan, FT

2014

## SKILLS

---

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

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

**Computer Vision:** Face Recognition, Video Analysis, Image Processing, OCR

**Others:** AWS EC2, Dataiku, Tableau, SAS, DAGSHub, Data Vision Control (DVC)

## REFERENCES

---

### Muhammad Zul Fadhli Bin Mohd Fauzi

Geoscientist, RiseHill Data Analysis

(+60)11 - 219 25248

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

### Syahirah Ehlam

Admin & Account, RiseHill Data Analysis

(+60)17 - 3018485

[syahirah.ehlam@rdagroups.com](mailto:syahirah.ehlam@rdagroups.com)