**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/)

# CAREER OBJECTIVE

# Strong analytical thinker with problem-solving and presentation skills. Seeking to leverage on model development experience and strong understanding in predictive modeling, and computer vision with a solid computer science background and understanding of machine learning & data science will be utilized. Bringing hands-on experience, excellent collaboration and communication skills, as well as programming skills.

# PROFESSIONAL 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)
* **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

**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++

**Shoppegram Cyberjaya, MY**

**Digital Marketer** 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 Kuala Lumpur, MY**

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

**Internship Kuching, SWK**

**Department of Information System** 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, CSS, SQL, & database management to complete the SL System & SL website within the time frame

**PROJECTS & COMPETITIONS**

**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 are 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>

**Integrated Data Intelligence (IDI)**

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

# EDUCATION

**UNIVERSITY OF MALAYSIA SARAWAK Kuching, SWK**

Bachelor of Computer Science with Honors (Information Systems)2019

**LABUAN MATRICULATION COLLEGE Labuan, FT**

Biology, Chemistry, Computer Science2014

# CERTIFICATIONS

[**DEEP LEARNING IN COMPUTER VISION**](https://cert.certifai.ai/8a4276ed-97e4-4d92-b2dd-fcfb7538a628?key=f4027d85f2161a4fb6efbe3f3d2116bd1ba62c0b02eaae2f1afe73dfe372dec1#gs.gzqfzm) **Skymind**

CERTIFAI30 April 2020 – 30 April 2022

**TAEKWONDO BLACK BELT 1st DAN Taekwondo Malaysia (TM)**

UNIMAS Taekwondo Club2018

# 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, 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

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**Muhammad Zul Fadhli Bin Mohd Fauzi**

Geoscientist, RiseHill Data Analysis

(+60)11 - 219 25248

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

**Syahirah Ehram**

Admin & Account, RiseHill Data Analysis

(+60)17 – 3018485

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