

HUSEIN ZOLKEPLI

Email: husein.zol05@gmail.com

Gitlab: <https://gitlab.com/huseinzol05>

Website: <https://huseinhouse.com>

Github: <https://github.com/huseinzol05>

Kaggle: <https://www.kaggle.com/huseinzol05>

Linkedin: <https://www.linkedin.com/in/husein-zolkepli/>

A software engineer focused on machine learning, devops, big data and data science. His hobby is chilling watching his flowers at his balcony.

Work Experience

Chief Technical Officer

Mesolitica Sdn Bhd, Kuala Lumpur, MY

December 2018 - Now

- Social media analysis and open source startup for Malaya, Natural-Language-Toolkit for Bahasa Malaysia, and massive Bahasa Malaysia corpus.
- Setup crawlers for Instagram, Facebook, local news, Twitter and Youtube.
- Setup massive deep-learning NLP engines infrastructure on top of Malaya, <https://malaya.readthedocs.io/>
- Setup Kafka cluster for realtime pubsub processing.
- Responsible to manage Kubernetes and Elasticsearch clusters and internships for our in-house development.

Data Engineer

Omnilytics Co, Shoppr Sdn Bhd, Kuala Lumpur, MY

August 2018 - November 2018

- Handling more than one billion unique fashion and beauty products across the world.
- Refactoring Elasticsearch indices to make it more efficient and less space used.
- Creating big data APIs for alternate sales index, sales forecasting, trend forecasting, market growth and market insight.
- Improving big data pipelining and scheduling using Luigi Spotify.

Data Scientist

XFERO Sdn Bhd, MaGIC, Cyberjaya, MY

February 2018 - October 2018

- Handled Dev Operations, task automations, quality assurance for branch merging and test units.
- Prototyped social media analysis dashboard, <http://www.huseinhouse.com/twitterdaily/>
- Developed deep learning integration models (computer vision, reinforcement learning) with drone DJI Tello.
- Implemented distributed data science jobs using Spark.

Executive Robotics (AI)

Robopreneur Sdn Bhd, Futurise, Cyberjaya, MY

March 2018 – August 2018

- Developed text to speech engine for small embedded devices, using MQTT and Websocket.
- Developed computer vision (face recognition, objection recognition, image semantic) engine for IOT devices, using MQTT and Websocket.
- Developed stacking attention Neural Machine Translation for humanoid chatbot (unpublished).
- Developed advanced face analysis (roll, pitch, yaw, emotion, gender, recognition, age, eyes open probability, mouth open probability).
- Developed Bahasa Malaysia speech-to-text engine and find similarity of entities related using pretrained local Bahasa Malaysia news.
- Developed parser Bahasa Malaysia natural language to chatbot language.

Machine Learning Engineer

DigitalHill Sdn Bhd, Kuala Lumpur, MY

August 2017 – July 2018

- Developed distributed sentiment Analysis engine supported Bahasa Malaysia and English for local news and social media on AWS.
- Forecasted social trends during General Election 15.
- Developed distributed intelligent crawlers.
- Developed Local Parliament text processing engine.
- Developed open-source bahasa corpus.

Open-source Developer

Devcon Community, MaGIC, Cyberjaya, MY

February 2017 – Now

- Developed open-source real-time deep learning computer vision and speech synthesis streaming.
- Conducted free classes related to deep learning, data science, data engineer, modern DevOps.
- Developed Vandalism Probabilistic Density API and integrated with Facebook Messenger.
- Maintaining Malaya Repository, Bahasa Malaysia NLP library with deep learning.
- Developed Bayesian and fuzzy engine to calculate probability density for KL traffic.

Founder

Soaning, Vega Residensi 1, Cyberjaya, MY

May 2017 – Feb 2018

- Developed stock engine that able to crawl selected hyper-parameters from Internet to predict specific stock.
- Developed API for local sentiment analysis, both overall and component-by-component study.
- Developed screening API for Github, able to classify developer personalities and technology concern.
- Developed screening API for CV / Resume, able to classify MBTI and developer personalities.
- Developed Twitter analysis API able study sentiment, emotion, polarity, irony, and subjectivity elements, correlation and boundaries.
- Developed news analysis API able study sentiment, emotion, polarity, irony, and subjectivity elements.
- Developed intelligent vision agriculture system able to study attractive attributes of flower and healthy percentage from crop.

Full Stack

Swift Enterprise, Kedah, MY

Sept 2016 – Feb 2017

- Developed Facebook Ads Management System using PHP, MySQL, JQuery, Bootstrap and ChartJS.
- Managed e-commerce system.

Open-source Projects

Malaya-Dataset

Text corpus for Bahasa Malaysia.

<https://github.com/huseinzol05/Malaya-Dataset>

Open sourced categorized articles, audience nationality, dependency parsing, 25k unique words dictionary, emotion, entities, gender, irony, karangan sekolah, language detection, categorized news, Twitter-based sentiment, news-based sentiment, multi domain sentiment, part of speech, polarity, political landscape, sarcastic based on news headline, stemmer, subjectivity, toxicity and subtitle.

Malaya

Natural Language Toolkit for Bahasa Malaysia, powered using deep learning models.

<https://github.com/huseinzol05/Malaya>

Supports news crawler, Entities Recognition, Language Detection, Normalizer, Num2Word, POS Recognition, Sentiment Analysis, Spelling Correction, Stemmer, Summarization, Topic Modelling, Topics and Influencers Analysis, Toxicity Analysis and Semantic Analysis.

Tello-Python

Tello packets prototype integrated with Tensorflow object and face detection models.

<https://github.com/DevconX/Tello-Python>

Reversed engineered Tello UDP streaming application to Python friendly and integrated with Tensorflow object and face detection during flight.

Deep-Learning-Tensorflow

Gathers Tensorflow deep learning projects.

<https://github.com/huseinzol05/Deep-Learning-Tensorflow>

Implementing various Convolutional Neural Network, Recurrent Neural Network, Feed-forward Neural Network, Sequence-to-Sequence, Hybrid, Bayesian optimization model, transfer learning, tensorboard, tf-distributed, tf-serving, tensorboard debugger, multiprocessing tfrecords, Regression and Generative Adversarial Network.

NLP-Models-Tensorflow

Gathers machine learning and tensorflow deep learning models for NLP problems.

<https://github.com/huseinzol05/Deep-Learning-Tensorflow>

Implementing text classification, chatbot, neural machine translation, embedded, POS-tagging, Entity-tagging, Question-Answers, Text-matching, Summarization, Attention model, Stemming, OCR, Language Detection, and Miscellaneous. Contain more than 220 deep learning models.

Stock-Prediction-Models

Gathers machine learning and deep learning models for Stock forecasting, included trading bots.

<https://github.com/huseinzol05/Stock-Prediction-Models>

Implementing stacking (machine and deep learning models) models, deep-learning models, trading agents, stock data analysis and fashion trending with cross-correlation. Covered LSTM, GRU, Sequence-to-Sequence, Attention model, and Differentiable Neural Computer.

Tensorflow-JS-Projects

Web projects using Tensorflow JS, Plotly, D3, Echarts, NumJS, and NumericJS

<https://github.com/huseinzol05/Tensorflow-JS-Projects>

Implementing tensorboard-like histogram on MNIST dataset, decision boundaries for IRIS using PCA and SVD, Linear, Polynomial, Ridge, Lasso and Elasticnet Regression, Stock forecasting and investment simulation with distribution study, Malaysia Export products forecasting and trading agent using Evolution Strategy.

Reinforcement-Learning-Agents

Gathers machine learning and deep learning models for Reinforcement Learning.

<https://github.com/huseinzol05/Reinforcement-Learning-Agents>

Implemented reward based for Evolution Strategy, policy gradient, q-learning, double q-learning, recurrent q-learning, double recurrent q-learning, dueling q-learning, dueling recurrent q-learning, double dueling recurrent q-learning, actor-critic, actor-critic dueling, actor-critic recurrent, actor-critic dueling recurrent and async q-learning.

Self-Driving-Car-Engines

Gathers signal processing, computer vision, machine learning and deep learning for self-driving car engines.

<https://github.com/huseinzol05/Self-Driving-Car-Engines>

Implemented Signal processing (1D smoothing, 2D smoothing, convolution 2 signals, pass-filters), Simple straight lane detection, Steering suggestion, Multi-lane detection, Multi-lane angle, Distance + Angle for object detection, Traffic light detection, Road Segmentation, Plate detection, Dynamic count lane detection and Gradient Smoothing.

Python-DevOps

Gathers Python stack for DevOps, everything is Docker!

<https://github.com/huseinzol05/Python-DevOps>

AutoPEP8, Pytest Flask, Flask, MongoDB, Flask Rest API, Postgres, Elastic Search, Luigi Spotify, Kibana, Distributed SocketIO with Redis, ELK Flask, Flask Hadoop, Mlfow Nginx, Flask Kafka, Flask Hive Hadoop, Pyspark Jupyter Hadoop.

Gather-Tensorflow-Serving

Gather how to deploy tensorflow models as much I can.

<https://github.com/huseinzol05/Gather-Tensorflow-Serving>

Implemented Object Detection using Flask SocketIO for WebRTC, Object Detection using Flask SocketIO for opencv, Speech streaming using Flask SocketIO, Multiple Inception with Flask using EC2 Docker Swarm + Nginx load balancer, Text classification using Hadoop streaming MapReduce, Text classification using Kafka and Text classification on Distributed TF using Flask + Gunicorn + Eventlet.

Machine-Learning-Data-Science-Reuse

Gathers machine learning and data science techniques for problem solving.

<https://github.com/huseinzol05/Machine-Learning-Data-Science-Reuse>

Implemented data preprocessing, natural language processing, suggestion engine, image processing, signal processing, stacking model, stochastic study, data visualization, markov, english-text normalization and sound processing.

Machine-Learning-Numpy

Code Machine learning models without any frameworks, Numpy only.

<https://github.com/huseinzol05/Machine-Learning-Numpy>

Implemented feed-forward, vanilla recurrent, lstm recurrent, gru recurrent, convolutional, batch-normalization, dropout, regularization, neuro-evolution, evolution strategy, clustering, decomposition, probabilistic, regression, trees based, timeseries and signal processing.

Neural-Network-Multilanguages

implement Artificial Neural Network on different languages.

<https://github.com/huseinzol05/Neural-Network-Multilanguages>

implement Gradient Descent Feed-forward and Recurrent Neural Network on different languages, only use vector / linear algebra library. Supported Julia, Python, Javascript, Go, C++, Ruby and PHP.

Pyspark-ML

Gathers data science and machine learning problem solving using PySpark and Hadoop.

<https://github.com/huseinzol05/Pyspark-ML>

Logistic regression, multinomial classification, topic modelling, word vector, principal component analysis, deep learning Tensorflow on Spark Cluster.

Online Data Science Competition

Zillow Prize: Zillow's Home Value Prediction (Zestimate) (top 19%, 748th of 3882)

Carvana Image Masking Challenge (top 86%, 623rd of 735)

Text Normalization Challenge - Russian Language (top 41%, 74th of 179)

Text Normalization Challenge - English Language (top 38%, 132nd of 426)

Porto Seguro's Safe Driver Prediction (top 8%, 284th of 3138)

House Prices: Advanced Regression Techniques (top 3%, 50th of 1934)

Titanic: Machine Learning from Disaster (top 1%, 53rd of 8649)

New York City Taxi Trip Duration (top 11%, 131st of 1257)

WSDM - KKBox's Churn Prediction Challenge (top 30%, 134th of 455)

Statoil/C-CORE Iceberg Classifier Challenge (top 19%, 632nd of 3390)

Corporación Favorita Grocery Sales Forecasting (top 19%, 308th of 1679)

Recruit Restaurant Visitor Forecasting (top 17%, 237th of 1432)

Hackathon

NASA Hackathon 2017

create a system to forecast earthquake.

GRAB Hackathon

create a system to predict traffic density.

Malaysia Airline Hackathon

create in-flight chat-bot system.

Advoc8 Hackathon

create a system to give multi-tag for pictures.

MYDD 2017

create a system to detect emotion and predict music based on the emotion.

Selangor Smart City 2017	create smart vandalism detection system.
MAMPU Data Hackathon 2017	create correlation study system.
TADHACK 2017	create crop and flower system.
FACE 2017	create smart house system.
REKATHON 2017	create in-house lidar system.
FisHackathon 2018	create fishermen forecasting system.
Makerthon 2018	main judge.
AWS Hackathon 2018	create Text Analysis for large company commerce dashboard.
MYDD 2018	create IOT tracking for turtles.
ZooHackathon 2018	create realtime animal detection dashboard.
Genting Programming	

Conduct Workshop Experience

Introduction to Data Science using Python

LHDN, Cyberjaya, Selangor

January 2019

- Numpy, Pandas, Matplotlib, Seaborn, Sklearn stacks
- Custom analysis on Malaysian taxpayers dataset
- Introduction to Tensorflow

Introduction to Python

BigIT, Kuala Lumpur

January 2019

- List, dictionary, comprehension, object oriented and lambda expression

Introduction to Data Science using Python

BigIT, Penang

January 2019

- Numpy, Pandas, Matplotlib, Seaborn, Sklearn stacks

Introduction to Docker-compose and Docker Swarm

Devcon volunteers, Malaysian Global Innovation & Creativity Centre

November 2018

- Config Flask and MongoDB using docker compose.
- Duplicating Flask with MongoDB using Docker Swarm on EC2 Docker machine.

Introduction to Computer Vision with Python

Undergraduate Computer Science, UiTM Tapah

October 2018

- How to use git, linux bash
- Introduction to Numpy and Matplotlib
- Implementing Convolutional Neural Network with Flask for serving application.

Self-driving car engine using Python

Public, Malaysian Global Innovation & Creativity Centre

September 2018

- 1D and 2D signal processing (blurring, filtering, convolution)
- Simple lane detection, straight lanes.

ELK Stack

Devcon volunteers, Malaysian Global Innovation & Creativity Centre

October 2018

- Elastic search using Python.
- Introduction to Kibana.

Self-driving car engine using Python

Public, Malaysian Global Innovation & Creativity Centre

September 2018

- 1D and 2D signal processing (blurring, filtering, convolution)
- Simple lane detection, straight lanes.
- Curve lane detection, polynomial lanes.

Introduction to Docker

Devcon volunteers, Malaysian Global Innovation & Creativity Centre

August 2018

- Explained VM, VE, Docker, LXC, Kubernetes, Docker Swarm
- Tutorial for custom images on Docker.

Data Science using Python

Software engineers, Brickfield College

May 2018

- Numpy tutorial.
- Pandas tutorial.
- Matplotlib, seaborn tutorial.
- Build our own suggestion engine using Nearest Neighbor algorithm.
- Pearson correlation study.
- Linear and non-linear regression.

Social Network Analysis

Public, Malaysian Global Innovation & Creativity Centre

April 2018

- Networkx tutorial.
- Social network visualization for Trump circle based on news and searching influencer similarity.
- Social network visualization for Najib Razak circle based on twitter and searching influencer similarity.

Internet of Things tutorial

Public, Malaysian Global Innovation & Creativity Centre

March 2018

- Setup ubuntu server on amazon aws EC2.
- Code socket-io server using node js.
- Code slave sensors using node js and python and visualize real-time data on remote dashboard.

Introduction to Deep Learning

Lecturers, UITM Jasin

February 2018

- Code feed-forward neural network using Tensorflow for Iris dataset on Google Colab.
- Code Convolutional neural network using Tensorflow for MNIST dataset on Google Colab.
- Code various gradient descent techniques using numpy.

Internet of Things tutorial

Public, Malaysian Global Innovation & Creativity Centre

January 2018

- Setup ubuntu server on amazon aws EC2.
- Code socket-io server using node js.
- Code slave sensors using node js and python and visualize real-time data on remote dashboard.

Speaker Experience

Sep 2018	NLP Bahasa Malaysia, Malaya, Devcon 5.
August 2018	NLP Bahasa Malaysia, Malaya, Pycon 2018.
June 2018	Real-time forecasting and analysis on front-end using Tensorflow.js
April 2018	Prevent botnet attack using big data, Universiti Pertahanan Negara Malaysia.
March 2018	Grill or Chill, Malaysian Global Innovation & Creativity Centre.
March 2018	Reinforcement learning for autonomous machines, PETRONAS Research Bangi.
March 2018	How to develop your own expert big data and IOT dashboard, UITM Shah Alam.
Feb 2018	Artificial Intelligence for game development, Malaysian Global Innovation & Creativity Centre.
Feb 2018	Introduction to Deep Learning, UITM Jasin.
Jan 2018	Emerging human emotion with Deep Learning, UITM Shah Alam.
Oct 2017	Data Science 101, Jetpack Kuala Lumpur.
Sep 2017	Vandalism Detection using AI Devcon 2, Malaysian Global Innovation & Creativity Centre.
Sep 2017	Vandalism and Theft Detection using AI, Cyberview.
Aug 2017	Life as a Machine Learning Engineer, ADAX Bangsar South.
July 2017	Intelligent Machine, Barcamp, MMU.

Academic Qualification

Sijil Pentaksiran Menengah MARA Junior Science College Taiping, Malaysia (2013)

Skill

Programming Language

C, C++, Java, Julia, Go, Ruby, Python, Javascript, PHP, R, C#

RDBMS

Postgres, MySQL, Oracle, RedShift, Hive

Apache big data

Hadoop, Kafka, Spark, Hive

NoSQL

Redis, MongoDB, DynamoDB, Elasticsearch

Cloud Service

Amazon, Google Cloud

Game Engine

Unity 3D (C# API), Unreal Engine 4 (Blueprint and C++ API)

Backend

PHP, Node JS, Python

ETL / Piping

Luigi spotify, Apache Airflow

Python big data

Dask, luigi, pyspark, hadoop streaming

Python data science / machine learning stack

Tensorflow, Matplotlib, seaborn, Scikit-learn, networkx, plotly, opencv

DevOps

Docker, Git, Kubernetes, Docker swarm, Docker compose, Gitlab CI, Circle CI, Travis, Jenkins

Language Proficiency

English: Good in writing and speaking.

Malay : Fluent in writing and speaking.

Referee

Referee 1:

DR. NORSHUHANI ZAMIN
University Malaysia of Computer Science and Engineering,
Precint 1, Putrajaya, 62000,
Malaysia
019-6626772
norshuhani@unimy.edu.my

Referee 2:

DR. AHMAD FAZREEN BAHARUDEN
DIGITAL HILL SDN BHD (833721-H)
Suite 3.2, Level 7, Putra World Trade Centre
41 Jln Tun Ismail, 50480 Kuala Lumpur, Malaysia
03-40444208
fazreen@digitalhill.com.my

Referee 3:

FAIZUL HIDAYAT BIN MOHD ARDANI
XFERO SDN BHD (1168264-H)
MaGIC, Cyberjaya, Malaysia
013-5848137