Raflie Zainuddin
rafolwen98@gmail.com | +60 18 278 4365 | linkedin.com/in/raflie

Data Scientist / Data Engineer

Data Science + Web Projects Portfolio: ye-yu.github.io/portfolio

Skillset Summary

- Machine Learning using sklearn & Keras – Python programming
- MapReduce, Hadoop, Spark – Java & Scala Programming
- RDBMS, Querying SQL and MongoDB
- Visualisation using Tableau & d3.js
- Other Programming languages: JavaScript, Kotlin, C
- Frameworks: Python Flask, JS Node.js, CSS Bootstrap
- Project Developments: Agile, Git, Github Actions

Recent Education

Bachelor of Computer Science (Hons) Majoring **Data Science** | Multimedia University, Malaysia
CGPA 3.89 (June, 2017 – June, 2020) – Scholar of Yayasan Telekom Malaysia

Internship Experience

Data Engineering Intern | MoneyLion Malaysia Sdn Bhd, KL | March – June 2019

- Programming ETL routine from **MongoDB** and other **SQL DB**s to **RedShift** using **Airflow** and **cron**
- Using Spark in **Java** and ***spark-submit*** to launch jobs and to extract 95% of GB-sized text files
- Configuring **spark** and **hadoop** in *nix system
- Designing web layout for internal program using **HTML + CSS**
- Using **Agile Project Management Framework** to complete 5 tickets every 2 weeks

Projects

[Utility App] July, 2020 (github.com/ye-yu/minecraft-data)
Minecraft Data – A game play data capturer

- Scope: Data collection, data engineering
- A client-side Minecraft mod for recording player play data using **Kotlin**
- Using a publisher thread, a consumer thread, and a shared ring buffer to periodically push player data
- Using byte encoding to reduce latency caused by fast periodic reading and writing

[Web App] June, 2020 (ye-yu.github.io/fast-read/)
Fast Read – An RSVP Utility App

- Scope: Web application and web design
- Read fast by flashing one word at a time in high speed using the client-side **JavaScript**
- Statically hosted app with features like momentum controlling, highlighting, and Chinese text support

[Data Science] June, 2019 – February, 2020 (github.com/ye-yu/cac-svs)
Extending Partial Singing Samples using Computer Assisted Composition and Singing Voice Synthesis

- Scope: Data science, unstructured data
- To extract audio features from singing voice input + To generate singing voice to complete input
- Training of Seq2Seq **neural network model** for note generation from singing songs
- Extracting timbre from voice as attribute for synthesising singing voice

Communication

Full Proficiency: **English** (MUET Score: Band 5), Bahasa Malaysia
Casual Speaking: Mandarin **Chinese**