CHEN-YI HUANG

(+886) 986 366 141
chenyihuang001@gmail.com
https://titaneric.github.io/
titaneric
chen-yi-huang

Persistent learner, dedicated graduate who major in data science. Skilled in data wrangling, machine learning knowledge and data visualization. Strong in Python and experienced with large-scale software code reading and tracing (E.g., CPython, PyTorch & Buddy System in Linux kernel).

Education

2018–2020 **Master of Data Science**, *Institute of Data Science & Engineering*, National Chiao Tung University, Hsinchu, Taiwan.

2014–2018 **Bachelor of Science**, *Department of Computer Science & Engineering*, Yuan Ze University, Taoyuan, Taiwan.

Master thesis

Title *Solving Traveling Salesman Problem with the Kernel-enabled Attention*

Supervisor Shi-Chun Tsai

Description We built on top of the prior state-of-the-art work who borrow the Transformer to solve the TSP. Motivated by the implicit dot product inside the kernel methods, we replace the scaled dot product with kernel in the attention mechanism. In our experiment, we archive shorter tour with a similar approach.

Work Experience

July 2018–Jan 2019 **Web Developer**, CS Computer Center, NCTU *Hsinchu, Taiwan*.

- Developed web services for **hundreds** of CS students, especially for Account Application System.
- Implemented in Laravel and Vue frameworks.
- Applied Git flow on development and Gitlab runner to automate test and deployment jobs.

Skills

Data Science Pandas, NumPy, PyTorch, Data Preprocessing, Data Visualization

Web Devs Laravel, Vue, MySQL, MongoDB, PHPUnit, Python unittest

System & Tools Linux, Git, Docker, Gitlab Runner, Travis CI

Programming Python (proficient), JavaScript, C, C++, Rust

Language Chinese (native), English (TOEIC 725)

Open Source Contributions

- Found the redundant calculation of derivative of power function in various deep learning frameworks. @ PyTorch, JAX, Autograd
- Developed some code to be more Pythonic. @ TensorFlow
- Bug reporting @ Python extension for Visual Studio Code
- Participation of issue discussion. @ Windows Subsystem for Linux

Projects

June 2020 **Real-time Traffic Anomaly Detector**.

- Anomaly detection in NCTU administration networks.
- Designed and implemented the data pre-processing pipeline with Apache Kafka, Spark and MongoDB.
- Processed data-stream in real time up to **30 kB/sec** and transformed into feature vectors to further predict.

Sep 2019 **Music Recommendation System**.

- Recommendation System using KKBox WSDM data
- Collected the additional data from Spotify and preprocessed them.
- Implemented the web-based interface to visualize the recommendation and user preference.

Sep 2019 **AutoDiff from Scratch**.

- Simple neural network library supporting auto-differentiation
- Reported an issue and solved it in various deep-learning libraries during the development.
- Enabled high-level layer usage and had already tested on real-world dataset.

Sep 2018 **Account Application System**.

- Web service for NCTU CS students to apply their account
- Developed in Laravel MVC architecture with additional Repository pattern.
- Designed the database schema.
- Implemented the business logic and account-activation-status API.

Certification & Award

- Arctic Code Vault Contributor @ GitHub
- Machine Learning with TensorFlow on Google Cloud Platform
- Querying Data with Transact-SQL