YIZE SUN

+852-90491349 | yizesun@connect.hku.hk
https://sun-yize.github.io/

EDUCATION

The University of Hong Kong Sep 2022 - Dec 2023

Master of Science - Artificial Intelligence

GPA： 3.54 / 4.3

Shandong University Sep 2018 - Jun 2022

Bachelor of Sciences - Statistics (Data Science & AI)

GPA： 3.70 / 4

Courses: Data Structure and Software Engineering (90), Operations Research (98), Deep Learning (96), Computer Vision (96)

PROFESSIONAL EXPERIENCE

Shanghai AI Laboratory Mar 2022 - Aug 2022

Python Developer Intern

- Worked on **SenseTime** dataset management platform project: **Sinan,** developed python client on dataset json schema validation and dataset search api, etc.
- Engaged in dataset standardization project: **OpenDataLab**, created standard json format for dataset annotation such as Box2D Tracking and Optical Flow, and developing python sdk for dataset standardization.
- Leveraged **OpenMMLab** and object tracking datasets such as ILSVRC, MOT to train models, and adjust these dataset annotation formats according to the model accuracy.

Cummins (China) Investment Co., Ltd Jul 2021 - Sep 2021

Data Analysis Intern

- Leveraged **Spark SQL** and **Python** to extract features from 2 million rows of raw data, including customer profiles, purchase history, inventory query history, etc.
- Built **Random Forest Regressor** using **PySpark** to predict the sales of mobile parts.
- Constructed **data pipeline** for weekly forecasting and accuracy calculation, achieving 30% average MAPE.

PROJECT EXPERIENCE

Smargo: An efficient and highly accurate solver for tsumego

- Built a Go game dataset: **Smargo Dataset**, which suitable for deep learning training, and solve the problems of insufficient tsumego data and inconsistent standards.
- Developed a Go game solver using Python, mainly using the **Monte Carlo Tree Search (MCTS)** algorithm and **multi-layer neural network** for chess position evaluation.

IndoorHIIT Motion Recognition

- Collected IndoorHIIT motion data from 50 testers, each data containing six-dimensional acceleration and angular velocity as well as number of motions; used **MySQL** database to select and preprocess data.
- Applied **Random Forest Classifier** for motion recognition; utilized **Wave Detection Method** to count the number of movements.
- Developed the interface of the WeChat mini-program; deployed the model in the server and used terminal cloud architecture to achieve motion recognition.

Object Detection and Image Classification Using Raspberry Pi

- Built **MobileNetV2** model to perform image recognition task based on 33,000 images of flowers and fruits using TensorFlow.
- Tuned the parameters of the model and applied batch normalization and data transformation to reduce validation error.
- Deployed real-time classification model on **Raspberry Pi using TensorFlow Lite**, achieving 90% accuracy.

HONORS & AWARDS

First Prize in Shandong Province, Contemporary Undergraduate Mathematical Contest in Modeling

Third Class Scholarship in 2020 and 2021 Academic Year

SKILLS LIST

- Programming Languages: Python, SQL, MATLAB, C
- Tools: Linux, Git, Docker, MySQL, AWS, K8s, Flask, Spark, Pandas, Numpy
- Languages: English, Mandarin