

YIZE SUN

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

EDUCATION

The University of Hong Kong

Sep 2022 - Jan 2024

M.Sc - Artificial Intelligence

GPA : 3.54 / 4.3

Courses: Statistics in artificial intelligence (A), Computational intelligence and Machine Learning (A)

Shandong University

Sep 2018 - Jun 2022

B.Sc - 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

The University of Hong Kong

May 2023 - Present

Research Assistant

- Conducted research on Motion Retargeting. Solved the problem of mismatch between pose detection and animation character skeleton. Used **Pytorch** and MediaPipe to conduct unsupervised training on a large number of online videos and built a motion retargeting model.
- Improved **3D Hand Tracking** algorithm. Used **deep learning** and kinematics method to calculate hand joint angle. Deployed the algorithm in the Unity application using **C++/C#** and MediaPipe.

Shanghai AI Laboratory

Mar 2022 - Aug 2022

Python Developer Intern

- Worked on SenseTime dataset management platform project: Sinan, developed **Python SDK** 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 **finetune DL 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

FortuneGPT: Financial Market Analysis Model Based on LLM and RL

- Developed the stock-picking model (FortuneGPT), integrating **Large Language Models** (LLMs) with inverse **reinforcement learning** to analyze Twitter data and stock market data, optimizing financial decision-making processes.
- Utilized **Python** to fine-tune the open-source LLaMa 2 on a **financial** text dataset, generating a GPT large language model specialized in the financial domain.

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, C/C++/C#, MATLAB, SQL
- Tools: Linux, Git, Docker, Pytorch, Tensorflow, Flask, Spark, MySQL, K8s, Unity
- Languages: English, Mandarin

孙易泽

电话: +86-17826486958 | 邮箱: yizesun@connect.hku.hk
个人网站: <https://sun-yize.github.io>



教育经历

- 香港大学 - 人工智能专业 硕士** 2022年09月 - 2024年01月
- GPA: 3.54 / 4.3
 - 人工智能统计 (A), 人工智能优化 (A), 计算智能与机器学习 (A)
- 山东大学 - 统计学 (数据科学方向) 本科** 2018年09月 - 2022年06月
- GPA: 3.70 / 4.0
 - 所在班级为校级实验班
 - 数据结构与软件工程 (90), 运筹学与数学建模 (98), 深度学习 (96), 计算机视觉 (96)

实习经历

- 香港大学 - 科研助理 计算机系** 2023年05月 - 至今
- 参与HKU Virtual Classroom项目, 该项目使用多个摄像头捕捉面部表情和身体动作, 构建 3D 虚拟教室和虚拟角色
 - 开展三维运动重定向研究, 解决了人体姿态估计与虚拟动画角色骨骼不匹配的问题, 运用Pytorch和MediaPipe对大量在线视频进行无监督训练, 并构建运动重定向模型
 - 改进的 3D 手部跟踪算法, 利用深度学习和运动学方法计算手部关节位置和角度, 并使用C++/C# 和 MediaPipe 在 Unity 应用程序中部署算法
- 上海人工智能实验室 - Python开发实习生** 2022年03月 - 2022年08月
- 开发商汤数据集管理平台 (司南) 的客户端, 使用Python开发客户端SDK的整体框架, 包括了数据集增删改查, 数据集格式校验、数据集检索等相关功能
 - 参与实验室数据集标准化项目 (OpenDataLab), 为目标跟踪、光流等数据集标注类型制定统一格式标准, 并开发用于数据集格式转换的Python SDK
 - 使用OpenMMLab训练目标检测和跟踪相关数据集, 用于标准化数据集的准确率验证, 按照训练结果和准确度进一步调整标准化数据集相关内容
- 康明斯 (中国) 投资有限公司 - 数据分析实习生** 2021年07月 - 2021年09月
- 使用Spark SQL从百万行原始数据中提取412维度的模型训练特征, 特征包括用户购买历史、库存及价格查询记录等方面
 - 通过Python对特征数据预处理, 并使用Pyspark的ML模块构建随机森林回归模型, 用于预测用户未来一周的零件购买量
 - 构建数据管道用于每周销量预测和准确度计算, 模型实际测试的误差MAPE降到30%以下

项目经历

- FortuneGPT: 基于LLM和RL的金融选股大模型**
- 开发了金融选股GPT (FortuneGPT), 将深度学习中的大语言模型 (LLM) 与强化学习相结合, 对twitter数据与股票数据进行分析, 以优化财务决策流程
 - 使用Python在金融文本数据集上对开源LLaMa 2进行微调, 并生成金融领域的GPT大语言模型
- IndoorHIIT动作识别项目**
- 采集50名测试者300余条健身数据, 每条数据包含个人信息以及运动的六轴数据, 将数据上传至服务器, 并使用MySQL数据库对数据进行收集和选取
 - 使用python机器学习算法进行动作识别, 通过随机森林模型和波峰检测法, 对测试者做出的动作进行实时识别和计数
 - 开发微信小程序界面, 同时将模型部署到服务器, 利用端管云架构实现实时动作识别
- 基于树莓派的花卉图像检测项目**
- 使用TensorFlow2构建MobileNetV2模型, 基于30种类别、共3万张的花卉图片进行图像分类模型训练
 - 调整模型参数及网络结构, 通过改变优化器和激活函数等方法, 提高模型精度
 - 利用TensorFlow Lite将深度学习模型部署在树莓派, 可以在移动端直接进行图像识别, 实际测试中准确率达到90%

荣誉奖项

- 2020、2021年度山东大学三等奖学金 2020年-2021年
全国大学生数学建模竞赛山东省一等奖 2020年9月

技能/其他

- 编程语言: Python, C/C++/C#, MATLAB, SQL
- 工具: Linux, Git, Docker, MySQL, AWS, K8s, Unity, Flask, Spark, Pandas, Numpy
- 语言: 英文、普通话