

# LIU Cheng

+ 86 15913169218 | chris.c.liu@outlook.com  
Personal Website: chris-chengliu.github.io  
Male | 172cm | 64kg

## EDUCATION

### Hong Kong Polytechnic University

Sep 2021

Master of Science in Data Science and Analytics, Faculty of Science

Hong Kong

- GPA: 3.35 / 4.30 (B+)
- Related courses: Deep Learning (A+), Artificial Intelligence Concept (A), Mathematical Modelling for Science and Technology (A-)

### Jinan University

Sep 2016

Bachelor of Engineering in Information Security, School of Intelligent Science and Engineering

Guangzhou

- GPA: 3.24 / 5.00 (82.5/100)
- Related courses: Advanced Mathematics (85), Probability & Mathematical Stat. (85), Linear Algebra (82), Math Model & Innovation (93)

## RESEARCH INTERESTS AND SKILLS

- **Interests:** Machine Learning, Computer Vision, Data Mining, AI with Medical Image, AI with Financial Data
- **Skills:** Linux, Python, PyTorch, MySQL
- **Languages:** English (IELTS 7.0), Cantonese

## PROJECT EXPERIENCE

### Credit Indicator Evaluation Model for Enterprise Business

Aug 2022 - Sep 2022

- Build an indicator system for enterprise business to evaluate the company's credit risk;
- Build an indicator pool by collecting data;
- Features are selected by using information value, correlation, and the importance in the LightGBM;
- Use LightGBM to do the classification and use grid search to find the suitable hyper-parameters.

### AI-dosimetrist: PolyU Student Entrepreneurial Proof-ofConcept (POC) Funding Scheme

Apr 2022 - Sep 2022

- Develop a model for the prediction of radiation dose based on CT images;
- Use Unet3D as an end-to-end network structure and Efficient-net as the backbone for feature extraction;
- Obtain a single-channel output image of the same size as the input for expressing the dose by modifying the network.
- The project was supported by the POC Fund of the Hong Kong Polytechnic University.

### Kaggle Google Smartphone Decimeter Challenge 2022

May 2022 - Jul 2022

- Develop a model based on the android smartphone GNSS chipset raw measurement data to improve the smartphone positioning resolution as much as possible;
- Parameter fusion model based on RTKLIB: gradually adjust key parameters on open source baseline to obtain offline validation results of training set;
- WLS (weighted least-squares estimation) model: smooth the data by carrier, then smooth the obtained coordinate positions and velocities using Kalman filter to obtain the prediction results.;
- Weighted fusion of the results of the above two models to obtain the final prediction results, which win a silver medal in the end of the competition.

### Kaggle UW-Madison GI Tract Image Segmentation

Apr 2022 - Jul 2022

- Develop a deep learning model to segment the tumor and gastrointestinal parts in MRI images.
- Using Unet as the segmentation model and Efficientnet-b1 as the feature extraction backbone network;
- Based on the requirements, Dice coefficient and 3D Hausdorff distance are used as the model evaluation metrics.
- The model in the public test set results up to 0.87025, in the private test set results up to 0.88416. The model win the bronze medal in the end of the competition;

### AI Automated Medical Image Analysis and Lesion Classification Model

Dec 2021 - Apr 2022

- Develop a system for analyzing liver CT images and detecting lesion types.
- Segmentation part: use U-Net on 3D liver CT image to do the segmentation and select the most representative slice;
- Classification part: use ResNet-50 to classify slices into 8 kinds of lesions, the accuracy rate is above 80%;

### **G-Research Crypto Forecasting**

Nov 2021 - Feb 2022

- Predict the return for the next three months based on data such as crypto transactions and bids;
- Construct a market index including 14 cryptos by extracting the features such as volume ratio of the last days;
- Building LightGBM and CatBoost models, and using RMSE as the loss function;
- Model ensemble: combine the results of the two model and get the final predict result, and the pearson correlation coefficients of between the predicted results and the real data on the test set is more than 0.01.

### **Classification of brain fMRI based on graph neural network**

Oct 2021 - Jan 2022

- To build a system for decoding and mapping the cognitive state of the human brain from fMRI images of the brain based on the HCP dataset.
- Use Graph U-Net structure to segment 3D brain fMRI slice images and select the most representative slices;
- Use GNN as the basic structure, the ROI in the brain area of interest is used as the network node, the pearson correlation coefficient is used as the node feature, and the partial correlation coefficient is used as the edge.

### **Personal Loan Default Forecast**

Sep 2021 - Dec 2021

- Build risk control models that capture the correlation between basic user information and default behavior in different businesses and help default prediction for new businesses;
- Analyzing the types of each attribute, extracting useful features, and constructing new features such as geolocation-related information;
- Use XGBoost and LightGBM to predict, and use finetune and migration learning on different target business training sets;
- Use cross-validation and stacking to do the model ensemble. The AUC of the model can achieve more than 0.9.

### **Optiver Realized Volatility Prediction**

Jun 2021 - Sep 2021

- To predict short-term volatility for the next three months based on data such as prices and bids.
- Extracting features such as bid-ask spreads, log returns, weighted average prices and real volatilities on the underlying data.
- Extracting features between similar stocks using K-means clustering; obtaining features at similar time points before and after the same stock using KNN classification.
- Training LightGBM and NN models on the basis of the above features, with RMSE as an evaluation metric less than 0.2.

### **Music Generation System Based on LSTM and GAN**

Feb 2020 - Jul 2020

- Develop a lightweight artificial intelligence composition system;
- Use MIDI files as the dataset, using the framework of Python and PyTorch, using LSTM network to learn the musical melody direction of the dataset, and using CGAN network to learn the accompaniment of the specified musical style;
- Conducting a blind questionnaire test, the result is that the music composed by the system is more than 80% similar to the music made by human;

### **Tramcar Environment Safety Monitoring System**

Sep 2019 - Jan 2020

- Developed a monitoring system for environmental safety and passenger safety in trams, including abnormal behavior of passengers, such as running and falling;
- Responsible for passenger abnormal behavior monitoring. The process of acquiring real-time images in the carriage and using C and Python language to build YOLO V3 Tiny pedestrian detection model, DeepSort target tracking model and ST-GCN action recognition model;
- 3 detection objectives of passenger detection, passenger trajectory tracking and passenger behavior detection are finally achieved;

### **Prediction of Residential Price Trend Based on Gray Model**

May 2018 - Apr 2019

- Explore the impact of various factors including policies on Hainan's housing prices, and make future commercial housing prices forecast;
- By comparing the influencing factors of Haikou and Sanya commercial housing prices in Hainan Province to establish the model.
- The gray model is used to predict the future trend of commercial housing prices and analyze their fluctuations.

### **An image processing platform based on AI**

Sep 2018 - Mar 2019

- The goal is to build a processing platform that combines basic image computing functions and machine learning algorithms.
- Using MFC and Python as development tools, it realizes basic operation functions such as picture addition multiplication point operation and interpolation-based image scaling, rotation and stretching, as well as image enhancement and edge detection in the frequency domain.

## TRAINING EXPERIENCE

---

### MicroMasters Program in Statistics and Data Science

Oct 2021 - Present

- Online training program offered by MITx;
- Master the foundations of data science, statistics, and machine learning.
- Analyze big data and make data-driven predictions through probabilistic modeling and statistical inference; identify and deploy appropriate modeling and methodologies in order to extract meaningful information for decision making.
- Develop and build machine learning algorithms to extract meaningful information from seemingly unstructured data; learn popular unsupervised learning methods, including clustering methodologies and supervised methods such as deep neural networks.

### Scalable Parallel Computing

Oct 2021 - Nov 2021

- Led by PolyU University Research Facility in Big Data Analytics
- Learn about parallel computing with RAPIDS
- Learn about parallel computing for machine learning using GPUs

### Machine Learning and Data Science

Jun 2020 - Sep 2020

- Led by Professor Mark Vogelsberger, MIT.
- The program covered the concepts in machine learning and data science with a focus on application development. Specifically, learned about the most common Python frameworks like Scikit-learn, TensorFlow, Keras, and PyTorch.
- Implemented a wide range of machine learning algorithms in these frameworks with a focus on deep learning, and the basics of big data processing with Hadoop and Apache Spark.

## WORK AND INTERNSHIP EXPERIENCE

---

### Bank of china limited shenzhen branch

Aug 2022 - Aug 2022

Intern, Department of Information Technology

Shenzhen

Learn the workflow of bank, common algorithms and models. Participate in the project of buliding a credit indicator evaluation model, and finally collate a comprehensive indicator system.

### Ericsson (Xi'an) information technology service co., ltd. guangzhou branch

Dec 2020 - Apr 2021

Integration engineer, EGX

Guangzhou

Learn the cloud platform construction and the use of OpenStack, K8s, Docker and other tools. Participate in the deployment and debugging of user plane function (UPF) module network element nodes.

## HONORS & AWARDS

---

Kaggle Google Smartphone Decimeter Challenge 2022 silver medal (17/571) 2022-07-29

Kaggle UW-Madison GI Tract Image Segmentation bronze medal (94/1548) 2022-07-14

Certificated by PolyU Research Facility in Big Data Analytics that successfully completed in Scalable Parallel Computing 2021-11-06

Certificated by MITx that successfully completed in 6.431x: Probability - The Science of Uncertainty and Data 2021-09-06

Certificated by MITx that successfully completed in 18.6501x: Fundamentals of Statistics 2021-09-06

Certificated by EdinburghX that successfully completed grade in PA4.1x\_MM: Introduction to Predictive Analytics using Python 2021-09-06

3rd Prize, the 12th National Contest of Electrical Mathematical Modeling 2019-07-06

1st Prize, the 9th MathorCup University Mathematical Modelling Challenge 2019-06-14

Meritorious Winner, 2019 MCM/ICM (Top 8% of the world) 2019-04-19

2nd Prize, the 1st China Geosciences Big Data Mining and AI Challenge 2019-04-06

## SELF-EVALUATION

---

- Have a strong interest in artificial intelligence products, have a good understanding of machine learning and its mathematical theory knowledge, familiar with commonly used models, algorithms and frameworks, such as linear regression, SVM, Tensorflow/PyTorch, etc.
- Familiar with Linux development environment and Python language, have the ability and experience to implement data mining and computer vision engineering.
- Excellent logical thinking ability, strong data analysis and problem solving ability, full of passion for solving challenging problems
- Self-motivated, full of enthusiasm for learning, good team spirit.