

# Yufei Ma

Email: [moana.yf@gmail.com](mailto:moana.yf@gmail.com) | LinkedIn: [linkedin.com/in/mayufei23/](https://www.linkedin.com/in/mayufei23/) | Phone: 857-208-1563

## EDUCATION

### The Chinese University of Hong Kong, Shenzhen

Shenzhen, China

B. S. in Data Science

Sep. 2019 – Jun. 2023

**Core Courses:** Calculus, Linear Algebra, Probability, Statistics, Optimization, Discrete Math | Data Structure and Algorithm, Database, Programming Paradigm | Machine Learning, Advanced Machine Learning, Deep Learning, Speech and Natural Language Processing, Data Mining.

### Boston University

Boston, MA

M. S. in Computer Information Systems

Sep. 2023 – May 2025

**Expected relevant courses:** Advanced Networking, Web Application Development, Software Design and Patterns, Mobile Application Development with Android System, Software Quality, Testing, and Security.

## PROFESSIONAL EXPERIENCE

### Bairong, Inc

Shenzhen, China

Machine Learning Engineer Intern

May – Aug. 2022

- Completed the project achieving the outcome of major clients' purchase with premiums. Here're my parts:
- Implemented an experimental Credit Scorecard Model with Logistic Regression in Python to show potential performance with customized data. We achieved high K-S score, and clients showed will to pay.
- Implemented a LightGBM algorithm with 3<sup>rd</sup> party data to enhance the models' performance. I also authored a script generating analytical reports to highlight the improvements.
- To better assess loaners' credibility, our team tried several ways of feature engineering. I generated 132 new features and 32 of them achieved remarkable Information Value.

### Ping An Asset Management Co., Ltd.

Shanghai, China

Financial Analyst Intern

Aug. – Sep. 2021

- Analyzed fixed-rate, fluctuating service fees, annual income indexes on 5-year asset contracts using SQL.
- Participated in releasing company's Q2 financial report by providing data cleansing support with Excel.

## COMPETITION

### Kaggle: Feedback Prize - Predicting Effective Arguments

Jul. 2022

- Implemented data preprocessing with NLP techniques like tokenizing, normalizing, and padding.
- Iterated pre-trained models like BERT, DeBERTa, and RoBERTa and parameters and got a silver medal.

## PROJECTS

### VOC2012 Object Detection Task - Deep Learning Course Project

May 2023

- Implemented traditional object detection algorithms with OpenCV with manual feature selection.
- Trained & fine-tuned STOA models including Faster R-CNN, SSD, and YOLO to compare performances.
- Wrote Bash scripts to train and log results after individually setting up cloud GPU resources.

### ChatGPT Detector - NLP Course Project

Mar. - May 2023

- Reproduced a RoBERTa ChatGPT/Human text classification model from a paper achieving >0.999 acc.
- Set up multiple LLM APIs and on-device models to produce more training text which achieved better generalizing performance over the baseline model.

### RLHF Project - Advanced Machine Learning Project

Apr. - May 2023

- Implemented a Chinese movie comment generating machine with certain given prompts using RLHF.
- Used ERNIE tokenizer and GPT-2 for generation. Reward model is based on sentiment analysis.
- The final generation is consistent after human ranking adjustment.

## ADDITIONAL INFORMATION

**Programming Languages:** Python, C++, Java, SQL

**Technologies:** NumPy, Dask, TensorFlow, PyTorch, JAX; Git, GitHub; CMAKE, Jenkins, Azure, GCP

**Language Skills:** Native Chinese speaker. Fluent in English

**Extracurricular:** CUHKSZ Dance Club: Director of the ballet department, principal dancer and main choreographer. Beijing Dancing Academy Ballet Grading Grade 8.