

# Yihan Zhu

+86 15601595757 | zhuyihan5023@outlook.com | [www.linkedin.com/in/yihan-zhu-1386b3217/](http://www.linkedin.com/in/yihan-zhu-1386b3217/)

## EDUCATION

<b>Columbia University</b> 4.0/4.0	New York, NY
MS in Electrical Engineering	Aug 2021 – Feb 2023
<b>Southeast University</b> 3.53/4.0 (86.21/100)	Nanjing, CN
BS in Electronic Science and Technology	Aug 2017 – Jun 2021

## RESEARCH EXPERIENCE

<b>Columbia University</b>	New York, NY
<b>Advanced Spoken Language Processing</b>	Sep – Dec 2022

- Extracted speech segments from the Emotional Prosody Speech and Transcripts, normalized the features accordingly by speaker, visualized the mean and standard deviation of each feature for all of the 15 emotion classes.
- Used the openSMILE toolkit to extract a set of acoustic-prosodic features, predicted the emotion classes by leave-one-speaker-out cross-validation.

<b>Algorithms of Determining DNA Health</b>	Jul – Aug 2022
---	----------------

- Deployed Aho Corasick algorithm to build up a Pattern Matching Machine (PMM) to achieve simultaneously keywords matching against target string in one pass through.
- Designed an improved Trie structure to simplify algorithm and achieved searching and comparison over 100,000 patterns strings, reduced complexity from  $O(N^2)$  which needs  $2 \cdot 10^{12}$  operations to  $O(N \log N)$ .

<b>Kernel Method for Deep Learning</b>	Feb – May 2022
--	----------------

- Investigated the geometric properties of arc-cosine kernels. Used arc-cosine kernel function family to simulate the computation of huge, multilayer neural networks based on MINST dataset.
- Designed different architectures for Multilayer Kernel Machines (MKMs) and tested the usage of kernel functions in shallow and deep structures.

<b>Decentralized Application on Blockchain</b>	Feb – May 2022
--	----------------

- Designed and created a decentralized GameFi which can adopt, feed, make friends and trade virtual patronus in Solidity and deployed it onto Ganache.
- Used MetaMask to access Ethereum wallet through the browser extension and exchange Non-fungible tokens (NFTs).
- Set up Frontend View Engine using Embedded JavaScript with HTML/CSS/JavaScript.

<b>FishNet: Deep High-Resolution Feature Representation on Small Datasets</b>	Aug – Dec 2021
---	----------------

- Implemented FishNet and ResNet network with the deep learning framework TensorFlow. Highlighted the significance of deep high-resolution representation for taking full advantages of limited datasets.
- Trained and compared a series of 99-, 150- and 201- layer FishNet models, a series of 16-, 44-, 56- and 160-layer ResNet models and Wide-ResNet16-4, 40-4 and 16-8 on the Imagenette dataset.
- Attained over 98% training accuracy and over 75% validation accuracy of all the three FishNet models after 20 epochs.

<b>Deep Reinforcement Learning based on Atari game</b>	Aug – Dec 2021
--	----------------

- Established Deep Q-learning model (DQN), Double Deep Q-learning model (DDQN), applied duel model over DQN and DDQN and adjusted the replay buffer size to improve stability.
- Trained on game PongNoFrameskip-v4 in OpenAI GYM, set up learning rate decays from 1 to 0.01 and batch size of 32. Achieved great performance of 21 scores on DQN and DDQN models on game Pong during test cases.
- Implemented parallel environment training to further speed up experiments and converged to the optima more rapidly, example runs show that only takes 5 minutes to train agents to play Atari Pong.

<b>Active Learning-Based ML Methods for Heart Disease Prediction</b>	Aug – Dec 2021
--	----------------

- Deployed and compared C4.5, SVM, Naïve Bayes, XG-Boost, Random Forest, and Neural Network methods.
- Achieved 95% of testing accuracy over heart disease classification based on dataset from Cleveland of UCI repository.
- Utilized MMC, Random, Adaptive multi-label active learning selection strategies to reduce the cost of labeling.

<b>Southeast University</b>	Nanjing, CN
<b>Personalized Movie Recommender System</b>	Feb – May 2021

*Joint International Research Laboratory of Information Display and Visualization*, Advisor: Prof. Chen Li

- Designed and improved a real time web crawler which crawls movie data from Douban website to build personal datasets.
- Set up relational databases (MySQL) to store and pre-process real data.
- Used NLP pre-trained models and RNN network to achieve emotion analyzation through movie reviews and precisely predicted trends of box office based on weighted different elements (type/director/actor).

## PROFESSIONAL EXPERIENCE

---

### Bitmain-Sophon

Shenzhen, CN

#### AI Algorithm Internship

Jun - Cur

- Led the team of developing AI micro server SE5 which bases on TPU accelerator. Built a cross compilation environment, transferred Deep Learning models based on deep model frameworks such as pytorch (Yolov5/7/8) to fp32 bmodels through NNTCtoolchain that can run on the platform and realize target detection, achieved int8 model quantization and adjusted the accuracy.

### Columbia University

New York, NY

#### Teaching Assistant

Sep - Dec 2022

- Providing assistance in the course “Virtual Net & Cloud Computing” and mentored 60+ students in master degree.
- Holding office hours for helping students in understanding class.
- Helping students with debugging projects of deploying applications/websites on cloud platforms and improving stability.

### Whale Cloud Computing Technology Co., Ltd

Nanjing, CN

#### Software Engineer Internship

Sep 2020 - Feb 2021

- Implemented a user information interaction web service, employed vertical sharding before sending messages to Online Charging System, applied message queue rabbitmq for asynchronous processing, stored various info to different databases: user basic info, integrations etc, to Redis, user addresses etc, to MySQL...
- Used the data locality property of big data processing framework such as MapReduce to achieve privacy preservation.
- Explored the usage of DL techniques and random forest algorithm of predicting the user' s fixed broadband off-grid behavior, ranked the importance of each variable that affects the fixed broadband quality.

### Tencent Cloud Computing Co., Ltd Cloud Service Platform

Nanjing, CN

#### Data Engineer Internship

Sep - Nov 2020

- Found out differences and mutual exclusion conditions of solutions of MySQL and selected the most suitable one considering characteristics of projects and clients.
- Combined the work with previous study, utilized virtualization to plentiful deploy test validation environments and high efficiently obtained self-test data.

## SKILLS

---

**Programming Language:** Python, Java, C/C++, HTML5/Css3/JavaScript, Dart, MATLAB, R, Solidity, Latex

**Development & Tools:** MySQL, MongoDB, neo4j, Redis, PyTorch, Tensorflow, Keras, Matplotlib, GCP, AWS, JProfile, PCB Design/Layout, FPGA, Praat

**Language:** Mandarin (Native Speaker), English (Fluent): GRE 333, Spanish (Medium)

## AWARDS

---

Columbia University Master of Science Award of Excellence

2023/05

Columbia University 2022 Spring Master Honors Student

2022/06