

Yanzhou Pan

8181 Fannin ST, Houston, TX, 77054

+1(832)873-1553 | yanzhoupan@outlook.com | LinkedIn: <https://www.linkedin.com/in/yanzhou-pan-008356186/>

EDUCATION

Rice University, Houston, TX

Master of Computer Science

Cumulative GPA: 3.7/4.0

Aug 2019 - Expected Dec 2020

Wuhan University, Wuhan, China

Bachelor of Computer Science

Cumulative GPA: 3.64/4.0 (top 10%)

Sep 2015- Jun 2019

UC Berkeley, Berkeley, CA

Exchange Student - Computer Science

Cumulative GPA: 3.6/4.0

Feb 2018- Jun 2018

PROFESSIONAL EXPERIENCE

LinkedIn Corporation, Sunnyvale, CA

Machine Learning Engineer Intern

May 2020-Aug 2020

- Raised the AUC value of job-search Click Through Rate (CTR) prediction by over 3.5%, using negative down sampling.
- Implemented deep learning APIs for job marketplace product with **TensorFlow**, including stochastic pooling and detail-preserving pooling, and finished the offline tests and online A/B tests for the new APIs.
- Designed and programmed a module analyzer for job posts at position one, saved 50 mins for every offline model test.

GeeTest, Wuhan, China

Machine Learning & Data Analysis Intern

Dec 2018-Jul 2019

- Developed 3 captcha rules for effective human-machine identification, including black-IP and mouse-tracking.
- Analyzed the captcha data based on **Pyspark** and **Hadoop**, found and stopped 5 groups of attackers who were trying to attack our clients' websites.
- Provided data analysis reports for over 15 clients and provided recommendations on preventing attackers.

Key Laboratory of Software Engineering, Wuhan, China

Research assistant

Feb 2018-May 2019

- Proposed a fast similarity measure for Time-Series data clustering, slashed the average running time by over 35%.
- Introduced a semi-supervised learning method called Constraint Propagation Approach (CPA), improved the clustering F-measure value by over 5%, with only a small amount of additional information.
- Programmed the proposed framework based on **KMEANS** and **DBSCAN** clustering algorithms.

PROJECTS

Hash Embedding for Facebook's DLRM Recommendation System

Aug 2020 - Present

- Proposed a novel hash-based embedding layer using **Locality Sensitive Hashing (LSH)**, the neural network can be trained with less than 50% of the total parameters, without affecting the AUC value.
- Built the **HashNet** with **Pytorch**, which made the neural network small enough to fit into a single GPU.

Qingzhi-Unmanned Restaurant Payment System

Jan 2017 – Oct 2018

- Developed an image recognition system based on Convolutional Neural Network (**CNN**) and **TensorFlow**.
- Achieved up to 97% accuracy in identifying daily dishes.

ACADEMIC ACTIVITIES

- **Publication:** Guoliang He, **Yanzhou Pan**, Xuewen Xia, Jinrong He, Rong Peng, and Neal N. Xiong, "A Fast Semi-Supervised Clustering Framework for Large-Scale Time Series Data", *IEEE Transactions on Systems, Man, and Cybernetics: Systems* PP(99):1-16 · August 2019, DOI: 10.1109/TSMC.2019.2931731 <https://ieeexplore.ieee.org/document/8807381>
- **Conference:** Conference Guest of the CCF (China Computer Federation) Big Data Conference Oct 2018

SKILLS

- **Programming:** Python, SQL, C, Java, Scala, Matlab, JavaScript, HTML, CSS
- **Tools:** TensorFlow, Pytorch, Pyspark, Hadoop, Angular, MXNET, SPSS, Oracle, MySQL