Resume: Jiachen Liu

7348814120 | amberljc@umich.edu
http://www-personal.umich.edu/~amberljc/
Objective: PhD Program in CSE

EDUCATION

University of Michigan *Sep 2018 - Jun 2020*

Bachelor of Science in Engineering in Data Science *Ann Arbor*

- **GPA:** 3.82/4.0 ; Minor in Mathematics
- **Coursework:** Data Mining and Machine learning (A), Deep Learning (Grad. A), Database System (A), Numerical Method (A+), Linear Regression (A), Computational Data Science (Grad. A)
- **Honors/Awards**: Dean's List (2018), Dean's List (2019), 2019 University Honors

Shanghai Jiao Tong University *Sep 2016 - Jun 2020*

Bachelor of Science in Electrical Computer Engineering *Shanghai, China*

- **Honors/Awards:** Best Technology Award in Design Expo (2017), Distinguished Academic Achievement Award, Dean's List (Top 5, 2017-2018), Undergraduate Scholarship (Top 30%, 2016-2018), Excellent Student Union Member (2017), Champion of the Freshman Cup of women's basketball in SJTU (2016)

Massachusetts Institute of Technology (Full-time visiting researcher in EECS CSAIL) *May 2019 - Sep 2019*

University of Navarra (Exchange Student) *Jun 2018 - Mar 2018*

RESEARCH EXPERIENCES

Computer Science & Artificial Intelligence Lab, MIT *May 2019 - Present*

Advisor: Professor Samuel Madden *Cambridge*

Project: High-dimensional Data Index: Product Quantization Under High Dimensional Statistics

- Proposed an ANN (appropriate nearest neighbor) search system based on product quantization (PQ) to support fast ANN searches in high-dimensional scalable dynamic databases, based on the unique properties of high-dimensional data.
- Optimized product quantization searching accuracy by adding distance estimation correction term, which comes from the property that vectors are nearly perpendicular to each other in high dimensional space.
- Optimized product quantization searching efficiency in two ways: (a) using spherical K-means to build coarse index to fit high-dimensional data's spherical distribution, and (b) using machine learning to identify the dominant features to prune redundant features in high-dimensional space.
- Developed two novel designs for product quantization to support index updating: (a) utilizing hierarchical clustering to replace K-Means while indexing and (b) utilizing the tree structure of hierarchical clustering to efficiently split or merge clusters to adapt index to the new data with different distributions.
- Implemented the proposed algorithm and state-of-the-art algorithms (LSH, HNSW, PQ) for ANN search tasks, with experiment results showing that our proposed model is 2 - 10 times faster than the baseline, while keeping high accuracy.
- Summarized the results in "Towards a High Dimensional Data Management System" and "Product Quantization Under High-dimensional Statistics", which is to be submitted.

Electrical Engineering and Computer Science, University of Michigan *Sep 2019 - Present*

Advisor: Professor Mosharaf Chowdhury *Ann Arbor, Michigan*

Project: Distributed Machine Learning Scheduling

- Proposing a 2-level cluster scheduler that can break Gang scheduling for ML jobs by scaling up/down number of workers to improve overall efficiency.
- Developing a reverse-resource offer model that the framework can give out flexible resources to the cluster scheduler for scheduling incoming jobs.
- Applying to AutoML jobs' scheduling in order to reduce job completion time and cluster utilization: 1) distributing the struggling ML configuration, 2) reordering the configurations based on their predicted runtime and 3) applying bin packing variant algorithm to schedule configurations in each stage
- Summarized the results of AutoML scheduling into paper, which is to be submitted to ICML2020

Electrical Engineering and Computer Science , University of Michigan *Feb 2019 - Apr 2019*

Advisor: Professor Honglak Lee

Project: Text-and-Vision-Fused Framework for Academic Paper Review

- Proposed a deep learning-based text-and-vision-fused framework, to distinguish academic papers of lower quality based on contents, vocabulary usage and image quality, aiming to efficiently pre-filter bad papers out of academic paper review.
- Developed a novel hierarchical LSTM structure combined with ResNet to enable the model to understand long text.
- Conducted experiments on the fusion model with the state-of-the-art model and demonstrated high accuracy on the test set.

Industry Operation Engineering Department, University of Michigan *Feb 2019 - May 2020*

Advisor: Professor Siqian Shen Sponsored by Didi Chuxing *Ann Arbor*

Project: Machine Learning Based Route Recommendation System

- Investigated machine learning and robust optimization to build a route recommendation system to predict a customized route choice for ride-sharing.
- Applied a deep neural network approach to cluster customers by mapping the historical data to a deep feature space, and tagging the clustered customers to make a personalized recommendation.

Publication

- **Product Quantization Under High Dimensional Statistics,** *In preparation*

Jiachen Liu, Lei Cao and Samuel Madden

- **Towards a High Dimensional Data Management System,** *In preparation*

Dong Deng, Lei Cao, **Jiachen Liu** and Runhui Wang

PROJECT EXPERIENCES

Fog Auto Collection Using Bionics Structure *May 2018 - Aug 2019*

Advisor: Roberto Dugnani (Participation in Research Program)

- Contributed to constructing a testing environment for an auto collecting device that can set the humidity, wind, temperature and sunlight levels using Raspberry Pi.
- Tested the auto collecting materials in different scenarios and tuned the surface structure to increase water adhesion.

Augmented Reality Ghost Hunting *Oct 2018 - Dec 2018*

Alternate Reality Initiative club

- Designed augmented reality games using Unity, aiming to beat a "ghost" enemy without being attacked.

Robust Real-time Object Detection *May 2017 - Aug 2017*

Advisor: Yu Zheng (SJTU Summer Design Expo)

- Implemented a system to support real-time object detection, which was robust to background noise, using SVM, the region-growing method and the HSV background removing method, which was shown robust to background noise.
- Realized real-time data transfers through Internet so that users could monitor the activities in certain area.
- Won the "Best Technology Award" in the Summer Design Expo at SJTU.

EXTRACURRICULAR EXPERIENCES

University of Michigan EECS *Sep 2019 - Present*

Instructional Aide EECS484 Database System *Ann Arbor, MI*

- Lead the weekly recitation classes and hold the office hours.
- Work as part of the Center for Learning and Teaching to support students' in the learning processes.

Student Union of Joint Institute *Jun 2017 - Aug 2018*

Vice President *Shanghai, China*

- Coordinated the work of the various departments under Student Union to ensure the division of labor and cooperation.
- Communicated with all departments' directors weekly to ensure reliable operation.
- Founded the debate team and directed the team to prepare for various competition.

COMPUTER SKILLS

C++ (Proficient), Python (Proficient), SQL (Proficient), R (Proficient), C, Matlab (Proficient)