Yuhao Zhang

yhzhang.info

Email : yuz870@eng.ucsd.edu
Mobile : +1-858-666-5299

Education

- **University of California, San Diego** — La Jolla, CA
  PhD in Computer Science; GPA: 3.77; Advisor: Prof. Arun Kumar — *Sept. 2017 – June. 2023(expected)*

  Courses taken: Advanced Compilers, Principles of Programing Languages, Database Systems, Advanced Algorithm, Machine Learning, Data Mining & Analytics, Advanced Data Analytics, Computer Vision, Introduction to Robotics

- **Nankai University** — Tianjin, China
  BS in Theoretical Physics — *Sept. 2012 – June. 2016*

Professional Experience

- **ADALab, University of California, San Diego** — La Jolla, CA
  PhD research, deep learning systems — *Sept. 2017 - Present*
  - **Panorama**:
    - Developed the first data system for tacking the unbounded vocabulary (labels that were not present during training) issue for video querying
    - Designed the system unified and domain-agnostic, and lets application users generalize to unbounded vocabularies in an out-of-the-box manner without tedious manual re-training of DNNs
    - Implemented the prototype with Python and TensorFlow. Tested with applications including face recognition, pedestrian re-identification, car model recognition, animal species recognition, etc. It can achieve between 2x to 20x higher efficiency and generalize well to unbounded vocabularies
  - **Cerebro**:
    - Developed Cerebro, a scalable system for high-throughput reproducible neural network model selection
    - Proposed a novel form of parallelism for ANN training called model hopper. This approach improves both throughput and reproducibility
    - In terms of resource efficiency, it can be the optimal choice over the TensorFlow parameter server, Horovod, or task parallelism. Built a prototype to support both TensorFlow and PyTorch

- **Greenplum R&D at VMware, Inc.** — Palo Diego, CA
  Software Engineer Intern, in-database deep learning — *Summer 2019*
  - Worked on incorporation of deep learning functionalities into Greenplum database, allowing training and inference of deep learning models with TensorFlow on database-resident data
  - Integrated my research idea, Cerebro, into the deep learning training infrastructure of Greenplum, boosting the efficiency (convergence speed) by 10x
  - Contributed to Apache MADlib open-source project. Developed the core feature of a major release

- **Opera Solutions** — San Diego, CA
  Data Scientist Intern, machine learning based scheduling system — *Summer 2018*
  - Tested and analyzed a production machine learning-based scheduling system for cinema theaters. Optimized the parallel programming model and IPC mechanisms, resulting in over 50% of reduction in execution time
  - Migrated the existing single-node applications to the Hadoop ecosystem for distributed computing. Redesigned the data source layer to reduce the data warehouse I/O overheads, with over 90% of improvement
  - Reduced communication overheads, increasing the capacity of the system from 80 parallel jobs to 200 $\sim$ 300 as the customer requested, concluding the 10-month-long project in 3 months

Publications

Y. Zhang and A. Kumar, *Panorama: A Data System for Unbounded Vocabulary Querying over Video*, PVLDB (2020)
S. Nakandala, **Y. Zhang**, and A. Kumar, *Cerebro: Efficient and Reproducible Model Selection on Deep Learning Systems*, ACM SIGMOD DEEM Workshop (2019)
Y.-H. Zhang and X.-Q. Li, *Three-generation neutrino oscillations in curved spacetime*, Nucl. Phys. B **911**, 563 (2016) [hep-ph/1606.05960]
C. Liu, Y. -G. Miao, Y. -M. Wu and **Y.-H. Zhang**, *Self-regular black holes quantized by means of an analogue to hydrogen atoms*, Adv. High Energy Phys. **2016**, 5982482 (2016) [hep-th/1511.04865]