Yuhao Zhang

yhzhang.info

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

Education

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

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

  Programming languages: Python, C++, SQL, Java, Scala

Professional Experience

- **ADALab, University of California, San Diego** — San Diego, CA
  PhD research, deep learning systems — *Sept. 2017 - Present*
  - **Panorama**:
    - Developed the first data system to tackle the bounded vocabulary and life-long learning issue in video analytics
    - Designed an end-to-end, domain-agnostic and efficient system, allowing the user to add new classes on-the-fly to their model without tedious manual re-training of deep neural nets
    - Implemented the prototype with TensorFlow in Python. 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
  - **Cerebro**:
    - Developed Cerebro, a resource-efficient distributed deep learning system
    - Proposed a novel form of parallelism for deep learning training called model hopper. This parallel paradigm improves resource-efficiency drastically
    - In terms of resource efficiency, it is the optimal choice over the TensorFlow parameter server, Horovod, or task parallelism tools like Python Dask, Vizier, or Ray. Built a prototype to support both TensorFlow and PyTorch in Python
    - Extended the system with various execution backends including Greenplum Database and Apache Spark in Python, Java, and C++
    - Presented the work at Spark+AI Summit 2020
- **Greenplum R&D at VMware, Inc.** — Palo Alto, CA
  Software Engineer Intern, in-database deep learning — *Summer 2019*
  - Worked on the first in-DBMS deep learning system, allowing training and inference of deep learning models with TensorFlow on database-resident data
  - Integrated my research project, Cerebro, into the deep learning training infrastructure of Greenplum Database, boosting the efficiency by over 10x
  - Contributed to Apache MADlib project in Python and SQL. Lead the development of a major release
  - This project has been incorporated into Greenplum and production-ready for VMware's customers
- **University of California, San Diego** — San Diego, CA
  Teaching Assistant of Course: Systems for Scalable Analytics — *Winter 2019*
  - TAed the course which is about large-scale data analytics with Big Data tools and machine learning
  - Developed the first edition of course assignments and auto-grading programs with Python and Bash. The assignments involve Python Dask, PySpark, AWS EC2/S3/EBS, and Kubernetes. These assignments have been adopted by the course ever since and used by 500+ students

Recent Publications

1. A. Kumar, S. Nakandala, **Y. Zhang**, S. Li, A. Gemawat, and K. Nagrecha Cerebro: A Layered Data Platform for Scalable Deep Learning, CIDR (2021)
2. S. Nakandala, **Y. Zhang**, and A. Kumar Cerebro: a data system for optimized deep learning model selection, PVLDB (2020)
3. **Y. Zhang** and A. Kumar, *Panorama: A Data System for Unbounded Vocabulary Querying over Video*, PVLDB (2020)
4. S. Nakandala, **Y. Zhang**, and A. Kumar, *Cerebro: Efficient and Reproducible Model Selection on Deep Learning Systems*, ACM SIGMOD DEEM Workshop (2019)