Xingyao Wang

https://www.linkedin.com/in/xingyao-wang

Email : xingyaow@umich.edu
Mobile : +1 734-272-2163 / +86 176-9577-6683

EDUCATION

- **University of Michigan** Ann Arbor, MI
  Bachelor of Science in Computer Science and Data Science; GPA: 4.00 / 4.00 (University Honors) *Sep. 2019 – May. 2022*
  - **Relevant Coursework**: EECS 492 Intro to Artificial Intelligence (2020 Fall), EECS 482 Intro to Operating System (2020 Fall), EECS 376 Foundations of Computer Science (2020 Fall), MATH 425 Intro to Probability (2020 Fall), EECS 281 Data Structure and Algorithms (Grade A+), EECS 370 Introduction to Computer Organization (Grade A+), EECS 280 Programming and Intro Data Structures (Grade A+, Top 1.36%), EECS 203 Discrete Mathematics (Grade A), LING 111 Intro to Language.
- **Tianjin University** Tianjin, China
  Bachelor of Science in Pharmacy and Computer Science; GPA: 3.89 / 4.00 (Top 4%) *Sep. 2017 – Jun. 2019*

WORK EXPERIENCE

- **Tencent** Shenzhen, China
  Software Engineer Intern (Deep Learning), Jizhi HPC Team, Technology Engineering Group (TEG) *May. 2020 - Sep. 2020*
  - **Honor - Star Employee of June in Exploration**: (1 out of 20) Awarded by Jizhi team for outstanding work in exploring and implementing up-to-date methods from literature to converge ResNet50 deep learning model under large batch (32k - 64k batchsize); also for improving model accuracy from 87% to 92% under large batch setting by actively combining AutoML with elastic GPU computing power (utilize idle GPU resources for interruptible hyper-parameter tuning purpose, used 512 Nvidia V100 concurrently at peak).
  - **Trained and Converged ResNet50 on ImageNet dataset in record-breaking 2 minute**: Train deep learning model ResNet50 on ImageNet dataset used to takes hours or even days to complete, while we break the world-record by doing it in 2 minutes 31 seconds on public cloud server and 2 minutes 2 seconds using RDMA. (Paper In Progress)
    1) Utilized Layer-wise Adaptive Rate Scaling (LARS) with novel dynamic trust ratio clipping for faster convergence, optimized LARS performance up by approx. 8% with parallelization of local learning rate calculations. 2) Improved the scaling efficiency of horovod-style distributed training to 91% (16 machines, 128 GPUs) and reduced network bottleneck by utilizing 2D-Torus allreduce, multiple NCCL streams, and Top-k gradient sparsification. 3) Implemented Deep Gradient Compression with momentum correction to aid accuracy under top-k gradient spasification. 4) Implemented AdaScale SGD from ICML'20 to cope with learning rate scaling problem on dynamic large batch. 5) Utilized parallel Bayesian Optimization and Population Based Training (PBT) algorithm to conduct hyperparameter search under manually specified learning rate schedule. 6) Experimented with techniques like auto-augmentation, ghost batch normalization, lookahead optimizer, final polishing, loss flooding etc to speed up convergence. 7) Implemented rectangle validation in TensorFlow for better evaluation of image classification models, converted ImageNet TFRecord dataset in sorted order for rectangle validation implementation.
  - **Mixed Precision Training and TensorFlow AMP Algorithm Optimization**: Improved performance for some deep learning models by increasing the number of tf operations that can be cast to FP16 for some models (e.g. increased 37% of nodes cast to FP16 in Transformer-Base model), which is achieved by help implementing and submitting "ProprageWhiteFwdThroughClearAndGray" function in Tensorflow Auto Mixed Precision algorithm to the TensorFlow community.
- **University of Michigan, School of Information** Ann Arbor, MI
  Undergraduate Researcher, Blablablab NLP Group *Sep. 2019 - Present*
  - **Multi-modal Tweet-GIF Interaction Understanding**: currently working on a project under the supervision of Professor David Jurgens to use deep learning models (BERT/EfficientNet) to modeling the relationship between a tweet (text) and a GIF (multi-frame images) replied to that tweet.
  - **Reddit Crosspost Automation Experiment**: currently working to use the Machine Learning approach to tackle the question: "given two random posts on Reddit, can we use an Machine Learning model to determine that they should be located in the same subreddit?"
  - **Data Crawling, Processing, and Annotation**: 1) Utilized Selenium, Scrapy, and Twitter/Reddit API to achieve automation in accessing web data. 2) Acquired and organized two million entry of GIF data using an SQL database. 3) Experienced in managing dataset annotation via google sheet and inter-rater agreement (Krippendorff alpha, Fleiss Kappa) calculation.

PUBLICATION

- **Lane Extraction and Quality Evaluation: A Hough Transform Based Approach**: by Xingyao Wang, Da Yan, et al. Proceedings of IEEE 3rd International Conference on Multimedia Information Processing and Retrieval, August 2020.

SKILLS AND CERTIFICATIONS

- **Skills**: Python (Proficient), C++ (Intermediate), TensorFlow (Intermediate), C (Intermediate), Horovod (Intermediate), RPC (Basic, built a file transfer tool with RPC), Docker (Basic), Git (Intermediate), SQL (Basic).
- **Coursera Certification**: Deep Learning Specialization; Python for Everybody Specialization.