Xinyu Chen

Machine Learning Engineer, Verint System Inc.

(505)267.5752 xinyu.chen@verint.com

14315 E 25th Ln, Spokane Valley, WA 99037

==========

PROFILE

I am a Ph.D student at Washington State University and a Machine Learning engineer at the Verint System Inc. After receiving a MS degree from the Computer Science Department of the University of New Mexico in 2019, I joined the HiPDAC lab to pursue the doctorate degree under the instruction of Dr. Dingwen Tao. I am interested in the field of Natural Language Processing, Parallel Computing and Graph Theories. I have strong background in software development on IBM mainframes due to my previous working experience. My research goal is to integrate the power of graph theories, parallel computing and machine learning techniques in searching for solutions of real world applications.

EDUCATION

2021	Washington State University	in progress
2014	University of New Mexico	M.S.
2011	Peking University	B.S.
1998	University of International Business and Economics	Certificate

RESEARCH EXPERIENCE

Graduate Student at HiPDAC lab, Washington State University

Memory Optimization for Parallel Influence Maximization Algorithm 8.2020– present

- Characterization memory footprint for parallel IMM algorithm
- Memory usage reduction with Huffman or bitmap coding and efficient queries

Machine Learning Engineer, Verint System Inc.

Self-supervised dialogue segmentation and summarization 5.2019 – present

- Use language models to detect dialogue boundaries with self-supervised approaches
- Parse dialogue structures for downstream natural language understanding tasks

Scalable Intent Mining on dialogue texts

- Classification of customer intents by fine-tuning language models
- Content clustering by mining k-NN graphs and contextual sentence embedding

Research Assistant at Data Science Lab, University of New Mexico

Distributed Learning and in-situ Analysis 1.2016 – 10.2019

- Parallel clustering for high dimensional numerical data on distributed sites
- GPU accelerated clustering method from random projection

Graduate Student in LANL's Data Science at Scale Summer School

Research on Using Thrill for HPC data analysis 5.2018

- Proof of concept of using Thrill library for Big-Data analysis of HPC simulations
- k-Means with Thrill library on molecular dynamics and plasma particle simulations

Research on Visualization of learning in deep convolutional networks 5.2017

- Visualization the training progress of convolutional neural networks
- Pruning convolutional kernels based on training visualization

Domain Decomposition in a Monte Carlo Supernova simulation 6.2016

- Develop a mini-app as proof of concept for the Improved Milagro algorithm
- Implement the Milagro algorithm into the real Monte Carlo simulation code

WORKING EXPERIENCE

Software Engineer at the Accounting Center of China Aviation (ACCA)

Fare-Audit Engine, Software Architect **6.2007 – 1.2014**

- Research on IATA(International Air Transport Association) auto pricing methods
- System design based on ATPCO(Air Tariff Publishing Company) standard

Maintenance on the SPA Auto Proration System, Software Developer **7.1998 – 8.2007**

- Maintain PL/I programs on IBM mainframe
- Develop Special Proration module for CA (Air China) and CZ (China Southern)

PUBLICATIONS

Conference Paper

- Xinyu Chen, Qiang Guan, Xin Liang, Li-Ta Lo, Simon Su, Zhengyong Ren, James Ahrens and Trilce Estrada, In situ TensorView: In situ Visualization of Convolutional Neural Networks. 2018 IEEE International Conference on BIG DATA, Seattle, USA, Dec 10-13, 2018.
- Xinyu Chen, Jeremy Benson, Trilce Estrada, KeyBin2: Distributed Clustering for Scalable and In-Situ Analysis, the 47th International Conference on Parallel Processing, Eugene, Oregon, USA, August 13 - 16 2018.
- Xinyu Chen, Jeremy Benson, Trilce Estrada, keybin Key-based Binning for Distributed Clustering, the 19th IEEE Cluster conference, Honolulu, USA, September 5 - 8 2017.
- Xinyu Chen, Trilce Estrada, Index Clustering: A map-reduce clustering approach using Numba, Proceedings of the 6th International Conference on Data Science, Technology and Applications, Madrid, Spain, Jul 24 - 26, 2017.

Workshop Paper

- Xinyu chen, Ian Beaver, Cynthia Freeman, Fine-Tuning Language Models For Semi-Supervised Text Mining, the 5th Workshop on Advances in High-Dimensional Big Data (AdHD), Virtual, Dec 10, 2020
- Xinyu Chen, Qiang Guan, Xin Liang, Li-Ta Lo, Simon Su, Trilce Estrada, and James Ahrens, TensorView: Visualizing the Training of Convolutional Neural Network Using Paraview. the First Workshop on Distributed Infrastructures for Deep Learning (DIDL), Las Vegas, USA, Dec 11, 2017.

Conference Poster

- Mariia Karabin, Supreeth Suresh, Xinyu Chen, Ivo Jimenez, Li-Ta Lo and Pascal Grosset, Using Thrill to Process Scientific Data on HPC, the 30th ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis, Dallas, USA, Nov 11 - 16, 2018.
- Xinyu Chen, Qiang Guan, Xin Liang, Li-Ta Lo, Trilce Estrada, and James Ahrens, TensorViz: Visualizing the Training of Convolutional Neural Network Using Paraview, the 29th ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis, Denver, USA, Nov 12 - 17, 2017.
- Xinyu Chen, David Huff, Platon Karpov, Ryan Wollaeger, Gabriel Rockefeller, Brendan Krueger, Exploring Performance of Domain Decomposition Strategies for Monte Carlo Radiation Transport, the 28th ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis, Salt Lake City, USA, Nov 13 - 18, 2016.