JINJIE ZHANG

Email: jayzhang0727@gmail.com ⋄ Phone: (619) 830-8031 ⋄ Website: link

EDUCATION

University of California, San Diego 09/2018-present
Ph.D. in Applied Mathematics. Advisors: Rayan Saab and Alexander Cloninger. GPA: 4.0

University of Chicago 09/2016-06/2018
M.S. in Statistics. Advisor: Lek-Heng Lim. GPA: 3.9

Beijing Jiaotong University (a.k.a. Northern Jiaotong University) 09/2012-06/2016
B.S. in Information and Computing Science. GPA: 3.8

WORK EXPERIENCE

Applied Scientist Intern, Amazon, Bellevue, WA 06/2021-09/2021
Flex Science Group

- Developed ensemble learning methods (e.g. bagging, boosting, stacking) for a bunch of time series forecasting models, (e.g. linear regression, ARIMA, exponential smoothing, and random forest), to predict daily package volume in specific stations.
- Explored and modified various deep learning models for time series forecasting based on RNNs, LSTM networks, GRUs, and Transformers (attention mechanism).
- Wrote SQL to extract, transform, and load data for analytic processing using Amazon Redshift and managed/ran Python packages of machine learning models via AWS Lambda and Amazon EC2.

Research Intern, Lenovo, Chicago, IL 01/2021-05/2021
Deep Learning and Algorithm Group

- Worked with research scientists from Lenovo Research to propose and improve low-light image enhancement algorithms via deep learning models, including GANs, autoencoders, and retinex based approachs.
- Developed a real-time object detection and segmentation system based on YOLO, U-Net and (Fast) R-CNN.
- Launched EC2 instances and deep learning AMI with TensorFlow/PyTorch (GPU) on Amazon Web Services (AWS) to build up and train our deep learning models.
- Ran TensorFlow Lite benchmark tools on Android and iOS devices and deployed our models to the product: selfie camera in a mobile device.

Graduate Student Researcher, UC San Diego, La Jolla, CA 2020/06-present
Deep Learning and Compressed Sensing Group

- Used TensorFlow 2.0 and Keras to modify geometric deep learning (GDL) frameworks on graphs including DeepWalk, GraphSAGE, and NeoDTI, which leads to a state of the art algorithm in machine learning.
- Unified active learning (AL) on graphs and graph convolutional networks (GCNs) to improve classification accuracy by 5% on citation network datasets: Cora, Citeseer, and Pubmed.
- Proposed a new active learning algorithm used to select coreset for semi-supervised learning and proved its generalization bound to unseen data, which outperforms previous algorithms and baselines.
- Designed a fast binary embedding algorithm for large scale image retrieval and tested its performance on Yelp and Flickr image datasets.

Research Intern, Mitsubishi Electric Research Laboratories, Cambridge, MA 2020/09-2020/12
Joint Research Project in Graph Signal Processing

- Focused on the generalization and stability analyses of graph scattering transforms (GSTs) that are competitive in a variety of graph classification tasks on protein data sets D&D and the scientific collaboration data set Collab.
- Developed a novel approach to prune the graph scattering transform to remove non-informative features on-the-fly and evaluated its performance in classifying 3D point clouds using ModelNet40 dataset.

Data Science Intern, Ernst & **Young, Beijing, China** 07/2018-09/2018

- Manipulated large customer data set based on big data platforms Hadoop, Spark to exceed business requirements.
- Performed statistical analysis using R and Python to predict and analyze the behavior of Alibaba users.
- Reduced the time from business data to business decisions using BI tools such as Tableau and Power BI.

PUBLICATIONS

Review full publication list via Google Scholar: profile link

1. J. Zhang, A. Cloninger, R. Saab, Sigma-Delta and Distributed Noise-Shaping Quantization Methods for Random Fourier Features. *arXiv preprint arXiv:2106.02614*, (2021).
2. J. Zhang, R. Saab, Faster Binary Embeddings for Preserving Euclidean Distances. *International Conference on Learning Representations (ICLR)*, (2021).
3. J. Zhang, L.-H. Lim, S. Friedland, Grothendieck Constant is Norm of Strassen Matrix Multiplication Tensor. *Numerische Mathematik*, **143**, 905-922 (2019).
4. S. Friedland, L.-H. Lim, J. Zhang, An elementary and unified proof of Grothendieck's inequality. *L'Enseignement mathématique*, **64**, 327-351 (2018).
5. J. Zhang, S. Zheng, On refined Hardy-Knopp type inequalities in Orlicz spaces and some related results. *Journal of Inequalities and Applications*, **169**, (2015).

SKILLS

1. 3 year experience in Amazon Web Services (AWS) ecosystem including elastic compute cloud (EC2), deep learning AMI, and SageMaker.
2. 4 years experience in CV, NLP, and deep learning frameworks including PyTorch, TensorFlow 2 and Keras.
3. 4 years experience in Python programming, data analysis and machine learning. Excellent in OpenCV, NumPy, SciPy, Pandas, Matplotlib, Scikit-learn and other scientific libraries.
4. Proficient in query language SQL and big data framework Spark.