YINWEI ZHANG

zhangyinwei1994@gmail.com · (+1) 520-486-7382 · YINWEI · https://yinwei-zhang.github.io/

Summary

Ph.D. candidate in Industrial Engineering. Masters in Statistics. 1.5 years of industrial experience in developing/deploying deep learning models. Solid theoretical background in machine learning and statistical methods. Proficiency in Python, data science libraries, deep learning frameworks, and SQL. Self-motivated individual with excellent communication and presentation skills.

Education

University of Arizona, United States

Ph.D. Candidate in Industrial Engineering — Aug, 2018 - May, 2023 (Expected)
Master of science in Statistics&Data Science — Aug, 2020 - Dec, 2022 (Expected)
Master of science in Engineering Management — Aug, 2016 - May, 2018

Jingchu University of Technology, China

Bachelor of science in Process Equipment and Control Engineering — Sep, 2012 - Jul, 2016

Related Coursework: Fundamental of Optimization, Experimental Design, Theory of Probability& Statistics, Statistical Machine Learning, Database Design in SQL, Project Management

Technical Proficiencies

- **Soft skills**: collaboration, communication, writing, leadership, presentation, quick-learner, data visualization
- **Programming skills**: Python, R, Matlab, SQL
- **Frameworks**: TensorFlow, PyTorch, Lua, Scikit-Learn, SciPy, Numpy, Pandas, Git, Linux, Docker, OpenCV
- **Awards/Certificates**: IEEM Honorable Mention Award, Roots2Resilience Scholarship, AWS Cloud Practitioner (on-going)

Experiences

Research scientist intern, ABB Ltd (Fortune 500), Raleigh, NC — Jan, 2020 - Aug, 2021

Highlights: rotated among research groups, went through the life-cycle of deep learning algorithms using PyTorch and TensorFlow

- Developed the pipeline for data preprocessing, model training, testing, and deployment for **image based regression and segmentation** tasks in a **multi-GPU Linux server**. Tools include **TensorFlow/PyTorch**, **TensorFlow Serving**, and **Docker**
- **Fine-tuned** the neural networks with different data fusion mechanisms, learning rate schedulers, optimizers, and backbone layers. Monitored the training and testing processes in **TensorBoard** and prevented overfitting by **early stopping**
- Worked with cross-functional teams to **analyze errors** of the existing pipeline. **Optimized** algorithms for data collection, and **proposed new algorithms** for data augmentation and result interpretations. The overall performance was improved by **30%**
- Cooperated with the project managers in identifying the key **evaluation matrics**. **Visualized** and **presented** the results to the technical leaders. The final achievements include **3 patents** and **2 reports**

Research Assistant, University of Arizona — Aug, 2018 - Dec,2022

Highlights: Proposed new methods for computer vision, signal analysis, and process monitoring tasks

- Proposed an **additive regression** method for image-based **anomaly detection** by incorporating **LASSO** and **Ridge** regularizations. Designed a gradient-based optimization algorithm that can run parallely for model estimation with a time complexity $O(n \log n)$ in MATLAB. The F1 score on the testing dataset was improved by 15%
- Developed a simulation pipeline in Docker by **Python and C++** for studying the failure of the perception system in the autonomous vehicle. Considered the **temporal correlation** of the failure events and train a parametric model by **EM algorithm** to predict future failures, improving the MAE by **40%**
- Detected bursts in the water distribution system by **multivariate process monitoring** methods such as MEWMA, MCUSUM, and T-chart in **R**. Trained the models based on the **spatial-temporal signals** collected from pipe sensors and determined the optimal parameters by **cross-validation**
- Proposed an **unsupervised object detection** method by LK optical flow, feature tracking, perspective transformation, and frame differencing using **OpenCV**. Boosted the results by postprocessing such as thresholding, connected component analysis, and morphological operations. The algorithm was transmitted to a Raspberry Pi mounted on an UAV for testing

Projects

Hand Disease Diagnosis by Temporal Data Clustering 2022

- Made detailed plans for data collection by repeatable medical experiments. Reduced the data dimensionality and extracted the features of the collected temporal data by the functional **principal component analysis**
- Selected the number of PC by the explained variance and clustered the PC scores via the **Gaussian mixture model**

Flight Delay Prediction 2022

- Explored and preprocessed the raw tabular data such as filling missing values and creating new features by **Pandas**
- Constructed a **Random Forest** regressor for delay prediction and searched for optimal parameters with **grid-search**. Interpreted the model by investigating the **feature importance** using **Scikit-learn** and **Matplotlib**

Presentation

- Xia, S., **Zhang, Y.**, Liu, J. (2020). Investigation of Curing Process Heterogeneity from Raman Spectrum via CP Decomposition. *INFORMS Annual Meeting*
- **Zhang, Y.** (2019). Tutorial: Applications of Spatial-Temporal Data Analytics in Industry. *Grand Lab Slam Workshop*, University of Arizona, Tucson
- **Zhang, Y.**, Liu, J., Lansey, K. (2019). Functional Data Analytics for Detecting Bursts in Water Distribution Systems. *INFORMS Annual Meeting*, Seattle
- **Zhang, Y.**, Liu, J., Son, Y. (2018). Effective and Efficient Moving Object Detection by a Moving Camera, *INFORMS Annual Meeting*, Phoenix

Publication

- Pan, F., **Zhang, Y.**, Liu, J., Head, L., Elli, M., & Alvarez, I. (2022). Quantifying error propagation in multi-stage perception system of autonomous vehicles via physics-based Simulation. *Winter Simulation Conference*, Singapore
- **Zhang, Y.**, Zhang, T., Liu, J., Kang, W., Liang, R., & Potter, B. (2022). Profile extraction for optical lens curing process with Image-based Regularized Tensor Decomposition. *Proceedings of the 2022 International Symposium on Flexible Automation*, Japan
- Nikravesh, Y., **Zhang, Y.**, Liu, J., & Frantziskonis, G.N (2022). A partition and microstructure-based method for large-scale topology optimization. *Mechanics of Materials*, Volume 166
- Peterson, R. L., Shea, K. D., Liu, J., Luque, K., Powell, J., **Zhang, Y**., Williams, D. K., Martin-Plank, L., Heasley, B. J., Phillips, L. R., & Crist, J. D. (2021). Family caregiving context: a pilot study. *The Arizona Nurse*, April
- **Zhang, Y.**, Lansey, K., & Liu, J. (2020). Detecting bursts in water distribution system via penalized functional decomposition. IEEM Conference (honorable mention award)
- Lee, S., Jain, S., **Zhang, Y.**, Liu, J., & Son, Y. (2020). A multi-paradigm simulation for the implementation of digital twins in surveillance application. IISE Conference
- Lee, S., Jain, S., Yuan, Y., **Zhang, Y.**, Yang, H., Liu, J., & Son, Y. (2019). Design and development of a DDDAMS-based border surveillance system via UVs and hybrid simulations. *Expert Systems With Applications*, 109-123

Patent

- System and method to generate augmented training data for neural network, No. PCT/US21/37798
- Robotic systems and methods used with installation of component parts
- Robotic systems and methods used to update training of a neural network based upon neural networks outputs

Others

President of INFORMS Student Chapter at the University of Arizona

- Wrote proposals and lead 4 members to hold social events, presentations, and recruiting events for graduate students

Teaching assistant for SIE 506, Quality Engineering

- Assisted students in understanding process monitoring methods such as hypothesis testing, CUSUM, and EWMA

Teaching assistant for SIE 533, Fundamentals of Data Science for Engineers

- Demonstrated the logistic regression, boosting, and decision tree using Python, Numpy, Pandas, and Scikit-learn