Zhangyue Shi

Tel: +1 (405) 762-2899 | Email: zhangyueshi97@gmail.com | Web: https://shizhangyue.github.io/

EDUCATION

Oklahoma State University (OSU) — Expected: Aug. 2023

Ph.D. in Industrial Engineering & Management, Minor in Statistics. GPA: 4.0/4.0 — *Stillwater, OK, USA*
- *Dissertation Topic: Advanced Data Analytics-based Quality Assurance for Smart Manufacturing*

Xi`an Jiaotong University (XJTU) — Jun. 2019

B.S. in Mechanical Engineering. GPA: 3.7/4.0 — *Xi`an, Shaanxi, China*

EXPERIENCE

Machine Learning Experience in Anomaly Detection

1. Online Process Monitoring in Additive Manufacturing, *Research Assistant*, OSU — Feb. 2020 – Nov. 2021
- Detected process anomaly during manufacturing via neural network (in TensorFlow), machine learning (in scikit-learn), and statistical method (in R), which could achieve 0.94 f-score in supervised monitoring and 85% true detection rate in unsupervised monitoring.
- Extracted low dimensional features from high-dimensional time-series data using neural network (LSTM-autoencoder).
- Applied supervised adaptive boosting monitoring; unsupervised one-class SVM and EWMA control chart monitoring from features.

2. Machine Learning Model Performance Enhancement in Distributed System, *Research Assistant*, OSU — Nov. 2021 – Present
- Enhance monitoring performance of neural networks in distributed manufacturing system via knowledge distillation (in PyTorch), which makes performance of convolutional neural network (CNN) at data-poor agent have a 0.05 increment to 0.8 in f-score.
- Distill useful knowledge from CNN at data-rich agent to improve CNN at data-poor agents while preserving data privacy.
- Collect time-series printing data for normal and anomaly condition at different agents and apply CNN to detect process anomaly.

3. Data Augmentation for Time-series Data in Manufacturing, *Research Assistant*, OSU — Aug. 2019 – Oct. 2021
- Developed a data augmentation method based on generative adversarial neural network (in TensorFlow) to generate high-quality data; improved classification performance of logistic regression by 0.01 to 0.88 in f-score compared with model without data augmentation.
- Proposed a filter layer in the generator to select augmented sample of high-quality based on time-regularized Hausdorff distance, which could discover underlying temporal relationship among the time-series data collected during manufacturing.

Other Machine Learning and Data Analytics Experience

1. Identification of Characteristics of Small Non-coding RNA with Aging, *Research Assistant*, OSU — Oct. 2021 – Present
- Find the correlation between small non-coding RNA with human aging via plasma and serum of healthy donors using machine learning methods, which could achieve 0.94 in R^2 and RMSE $\leq$ 3.7 in test set based on the established age prediction model.
- Employ differential expression and conduct correlation analysis (maximum information coefficient) from transfer RNAs (tRNA), microRNAs (miRNA), etc.; and utilize adaptive boosting, gradient boosting, and random forest to build up regression model.

2. Surface-related Machine Parameter Identification in AM, *Research Assistant*, OSU — Aug. 2019 – Jul. 2022
- Identified relationship between machine parameters and resulting AM surface morphology via neural networks, machine learning and statistics, achieving 0.75 f-score in a 16-class classification, which could be further used in parameter selection and anomaly detection.
- Extracted low-dimensional features from high-dimensional surface data with heavy noise via neural network (Robust autoencoder).
- Together with conventional statistical feature, correlated extracted features with machine parameters via random forest.

3. Blockchain-enabled Data Storage in Manufacturing System, *Research Assistant*, OSU — Apr. 2020 – Jul. 2021
- Proposed a data storage method based on blockchain, which could detect malicious modifications on data via mismatch of hash value.
- Applied asymmetry encryption method and a camouflage method on data (in Python) to reduce risk of unauthorized access on data.

SKILL

Programming: Python, R, MATLAB, SQL — **DS Library**: Scikit-learn, PyTorch, TensorFlow, Gurobi, Pandas, Numpy

Hardware: Raspberry Pi — **Language**: Mandarin (native), English (fluent), German (basic)

Quantitative Methodologies: Machine Learning, Deep Learning, Statistical Inference, Knowledge Distillation, Continuous Learning, Model Compression, Data Augmentation, Statistical Process Control, Linear Programming, Time Series Analysis, Stochastic Process.

SELECTED HONORS AND AWARDS

- **Robberson Research and Creative Activity Grant**, The Graduate College, Oklahoma State University, 2022.
- **Outstanding Graduate Student**, The College of Engineering, Architecture and Technology, Oklahoma State University, 2022.
- **National Science Foundation (NSF) Student Support Award**, North American Manufacturing Research Conference 49, 2021.
- **Data Challenge Competition Finalist**, Quality, Statistics and Reliability (QSR) Section, INFORMS Annual Meeting, 2019.
- **First Prize**, National College Students Mathematical Modeling Competition Undergraduate Group Shaanxi Division, China, 2017.

SELECTED PUBLICATIONS

- **Shi, Z.**, Mamun, A. A., Kan, C., Tian, W., & Liu, C. (2022). *An LSTM-autoencoder based online side channel monitoring approach for cyber-physical attack detection in additive manufacturing*. **Journal of Intelligent Manufacturing**.
- 7 published and accepted in total, 6 first-author papers. A full list of publications can be found here.

COMMUNITY ENGAGEMENT

President, INFORMS Student Chapter at OSU | **Reviewer**, Journal of Intelligent Manufacturing, Healthcare Analytics | **Member**, IISE, INFORMS