Bo Xu

b68xu@uwaterloo.ca | 1-226-978-8738 | bo-xu-uwaterloo

EDUCATION

University of Waterloo, Canada **Sep 2019 – Present**

PhD. in Control Science and Engineering with Machine Learning (GPA: 89.14)

Beijing University of Chemical Technology **Sep 2015 – Jun 2019**

Bachelor of Science in Chemical Engineering and Technology

SKILLS

Software Development: Python, C++, R, Java, SQL, HTML, JavaScript, Git, MATLAB

Library: PyTorch, TensorFlow, ETL, Spark, Pandas, NumPy

PROFESSIONAL EXPERIENCES

Anaergia Inc. **Oct 2021 – Apr 2023**

Machine Learning Engineer & Data Scientist Intern

Optimization of Waste-to-Resource Plants with ML Prediction

- Undertook a comprehensive study of waste-to-resource plant processes for operational data cleaning, identifying key process and quality variables by feature selection and consult chemical engineers with domain, thereby enhancing overall operational understanding.
- Examined and preprocessed over a million data points using tools like NumPy, Pandas, and Spark ETL jobs for data imputation, in which I used a special KNN imputation method, reducing data anomalies by 15% and elevating the data quality for subsequent analysis.
- Employed advanced machine learning algorithms like Random Forest Regression, RNN, LSTM, and My own Dynamic methods. Evaluation matrices such as MSE, R^2, RMSE, and fitted plot were used to enable the product to reach its target goal with an error rate below 5%.

Anomaly Detection Strategy for Imbalance Data

- Designed a robust process monitoring and diagnosis framework for the imbalance data in Membrane Bioreactor process.
- Developed a special classification method for imbalance data. First, two statistical indices were designed to classify normal and abnormal samples where the limits of indices were calculated by normal samples. Then Random Forest Classifier was used to classify abnormal samples.
- Evaluated the results by Confusion Matrix, Recall, Precision, F1 score and Accuracy.
- Enhanced process oversight and diagnostic speed by 30%.

Multi-Objective Reinforcement Learning

- Devised a Reinforcement Learning model for the membrane bioreactor plant, creating an automated operation plan that generated Biogas with the target composition ratio. This innovation resulted in a 5% increase in operational efficiency and optimized decision-making processes.

AI Lab in UW **Sep 2019 – Present**

Research Assistant

Dynamic Weighted Partial Least Squares Algorithm

- Developed a novel dynamic weighted partial least squares (DWPLS) algorithm to solve redundant information in high sampling speed time series to enhance the prediction result. Initiated a project focused on improving dynamic data modeling; Compared Gaussian, Bayesian, and polynomial basis functions to enhance the dynamic-inner PLS model on prediction accuracy, convergence rate and time lags.
- Applied the model on 250,000 chemical lab operational data to simulate fault detection monitoring system, the new algorithm led to a 5% increase in fault detection rates than the baseline.
- The proposed method is shown to be able to track changes in process conditions in real time and identify potential quality issues. Applied the model with real industrial data and resulted 95% accuracy.

Dynamic Auto-regressive Latent Variable Model

- Developed a dynamic autoregressive latent variable algorithm that innovatively utilized both historical information of X and Y to predict current Y, enhancing convergence rate by reducing 10% training time and prediction accuracy improved around 13% on the high dimensional chemical operation dataset.
- Introduced an efficient dynamic autoregressive canonical correlation analysis (EDACCA) with economic SVD, which improved the model's computational efficiency and convergence rate.
- Addressed missing data issues by developing a robust data interpolation method with the dynamic model to fill in gaps caused by irregular sampling rates, leading to a 20% decrease in MSE.

PUBLICATIONS

My Google Scholar: https://scholar.google.com/citations?user=8OaLpBAAAAAJ&hl=en