YU YANG

Mobile: (+1)612-391-5600 Email: yang6367@umn.edu Webpage: https://yuyangyy.com

EDUCATION

University of Minnesota, Twin Cities *August 2018 - May 2023 (Expected)*
Ph.D. student in Statistics GPA: 3.978/4 Minneapolis, MN

Shanghai University of Finance and Economics *September 2014 - June 2018*
Bachelor of Science in Statistics GPA: 3.89/4 Shanghai, China

EXPERIENCE

Seagate Technology 09/2019–Present
Research Assistant (Advisor: Prof. Xiaotong Shen; Mentor: Sthitie Bom) *Minneapolis, MN*

Critical Sensor Selection for Abnormality Detection 09/2019–06/2020

- Proposed a multicollinearity-adjusted feature selection procedure for identifying critical sensors
- Constructed interpretable classification models to automatize abnormality identification
- Proposed several classification confidence measure for imbalanced data scenario

Causal Structure Learning in Manufacturing 06/2020–Present

- Explored large-scale and multi-sourced datasets and proposed systematic preprocessing pipelines
- Proposed causal structure learning methods to unveil the causal relations among abnormal events
- Proposed evaluation methods to examine the quality of the learned network without ground truth

PROJECTS

Causal Discovery for Mixed Data with Temporal and Group Constraints 01/2021–04/2021

- Proposed three causal discovery methods for high-dimensional mixed data with special constraints
- Performed experiments on the simulated data to examine the performance of the three methods

Topic-Aware Abstractive Text Summarization 01/2021–04/2021

- Proposed a new model by marrying Pointer-Generator Networks with Replicated Softmax RBM
- Experimented the model on the CNN/Daily Mail data

Retro-BiDAF: A Retrospective Reader Over BiDAF 10/2020–12/2020

- Proposed a question answering model for the SQuAD 2.0 Challenge
- Examined the idea of retrospective reading in the non-PCE scenario

Kaggle: Lyft Motion Prediction for Autonomous Vehicles 09/2020–11/2020
*Team Members: Xuesong Hou, Chunlin Li, Yu Yang (Ranked **top 6%**)*

- Explored the large-scale image data and visualized the paths of vehicles
- Built an ensemble model upon EfficientNet and DenseNet to predict the motion of on-road objects

Wells Fargo Campus Analytics Challenge 2020 07/2020–08/2020
*Team Members: Xuesong Hou, Chunlin Li, Yu Yang (Won the **Grand Prize** of the year)*

- Identified proper encoding schemes from model fitting details and proposed a top-performing classifier
- Proposed a novel method called Sparse Grouping Pursuit which efficiently reduced feature dimensions

R Package: ImbCalib – Probability Calibration for Imbalanced Data 04/2020–05/2020

- Wrote an R package to calibrate probabilities for imbalanced data
- Compared probability calibrations visually and quantitatively

MinneMUDAC 2019 Student Data Science Challenge 09/2019–11/2019
*Team: Women in Math and Stats (Won the **Analytical Acumen Award**)*

- Collected data from a wide range of sources and applied creative feature engineering
- Built an ensemble model upon XGBoost, LSTM, and VAR to predict the soybean futures closing prices

Learning Rate Decaying Scheme Investigation 11/2019–12/2019
Team Members: Liwei Huang, Yu Yang

- Proposed several learning rate decaying schemes and applied them to MNIST and CIFAR-10
- Analyzed the decaying schemes in terms of convergence time and model performances

Kaggle: Travelers Claim Fraud Detection 11/2018–12/2018
*Team Members: Somyi Baek, Sam Piehl, King Yiu Suen, Xun Xian, Yu Yang (Won the **2nd** place)*

- Proposed a new feature which greatly improved the predictive capability
- Constructed an ensemble model for prediction and applied LIME for interpretation

TECHNICAL STRENGTHS

Languages	Python, R, Shell Scripting, C/C++, SQL, HTML
Tools	Git, VS Code, Google Cloud, AWS, LaTeX