Shuai Wang | Curriculum Vitae

Cincinnati, Ohio 45230

(937)-654-0806 • vanstark88 at gmail.com • shuaiwang88.github.io

Education

Wright State University **Dayton, Ohio**
Ph.D. in Engineering Program, Industrial and Human System *2011–2017*

Dalian Jiaotong University **Dalian, China**
Bachelor of Management, Supply Chain Management *2007–2011*

Computer Skills

Language: Python, R, Julia; SQL

Machine Learning: regression, classification, clustering; xgboost, lightgbm, adaboost, random forest, etc

AutoML: Datarobot, H2O, Knime; pycaret, caret

Optimization Modeling: AIMMS, Pyomo, Minizinc, JuMP, OPL; Cplex, Gurobi, CBC, LocalSolver

Heuristic: Simulated Annealing, Genetic Algorithm, Tabu Search, Golden Section Search, etc.

Time Series: Anomaly detection: KATS; Forecasting: ARIMA, Prophet, RNN

Discrete Event Simulation: Arena, AnyLogic, Sigma

Database: Snowflake, Redshift, Clickhouse, Postgres, MySql; dbt

Cloud: AWS Lambda, Sagemaker, Spark, Airflow, API Gateway, EC2

Visualization: Shiny, Looker, Superset, Tableau

Experience

Pacific Life Insurance **Newport Beach, CA**
Principal Data Scientist *2023.6–Current*

Thomasnet **New York, NY**
Principal Data Scientist *2021.4–2023.5*

Reporting to head of data and AI. I started several 0 to 1 data initiatives at Thomasnet including:

- **Internal facing:**
  1. Applying AutoML to models like customer churn, bot detection, and sales forecasting.
  2. Recommending products portfolio to suppliers to maximize ROI using integer programming optimization.
  3. Creating data cleaning, extraction, normalization python library for all the enterprise data ingestion pipeline.
  4. Introducing new tools like Superset, Datahub, and Amundsen to the enterprise for better data governance and visualization.
- **External facing:**
  1. Creating time series anomaly detection (signals) of sourcing activities for institutional investors.
  2. Creating sourcing trend and stock price/revenue correlation for institutional investors.
  3. Working on TMX index to track supply chain status in the USA. TMX is used by researchers from Carnegie Mellon University for Congressional hearing on supply chain, and University of Hong Kong for COVID19 related studies.
  4. Creating APIs and charts for the Thomasnet's alternative data platform for institutional investors.

Kroger/84.51 **Cincinnati, OH**
Lead Operations Research Scientist *2017.9–2021.4*
I transferred with Kroger's R&D team, to Kroger's subsidiary data analytical company 84.51 in February 2018.

- **Order Forecast and Picking Staff Daily Scheduling System Optimization**: The daily staff scheduling system is built to optimize the number of staff required to picking orders at each hour. Time-series based machine learning model is implemented to predict the orders. An optimization model using CBC/Cplex solver is created to minimize labor cost. The project saved about 20% to 30% labor cost ($200 to $300 million) than the previous implementation. This project was escalated as one of top priority projects responding to COVID-19 at Kroger.
- **In-store Inventory Control and Restock Optimization**: The project is to transform the in-store inventory into scientific based management system, to increase productivity: 1. I build the inventory replenishment routing model, to reduce travel distance, using heuristics based on traveling salesman problem. A novel distance metric was proposed to match different store layouts. 2. A comprehensive staff productivity graph was created to track the scanning activity, idle time, travel time. 3. Heuristic-based restock strategy was created to alert the restock point based on BOH.

Large Scale Optimization Lab(Lab 249) from WSU partner with Kroger **Cincinnati, OH**
Operations Research Consultant at Kroger *2012–2017*

- **Product Promotion Planning and Assortment Optimization**: A MILP optimization model (Collaborative Category Optimization) was jointly developed with A.T. Kearney using AIMMS to build promotion planning and assortment selection to maximize the adjusted gross revenue. The system is implemented in 2014 and saves about 3% to 5% ($2 to $5 billion) of the total revenue annually.
- **Forecast and Optimization for the Little Clinics and Nurse Scheduling**: The problem is to accurately forecast the number of patients by type to each clinic. Sophisticated forecast models that utilize inputs from time, weather, social media data like Google trend queries are used to predict the number of visits. These forecasts serve as input to calculate the inventory necessary for each type of sickness, and the allocation of medical personnel and their shifts. The overall goal is to improve customer service and increase the number of clinics from 136 to 500 in three to five years.
- **High Value Pharmacy Product Local Inventory Transfer**: The problem is to ship medical drugs from stores with excess to stores with needs so as to: 1) reduce potential obsolesces in inventory; 2) better position drugs to meet customer demand; 3) to aggregate the shipment in such that transportation costs are reduced. The model is piloted in 121 stores and is expected to produce $30 to $50 million savings as well as reduction in out-of-stock.
- **Periodic Vehicle Routing**: Each Kroger store has demand that fluctuates within a week in various categories, such as frozen, fresh, and grocery. Determining the frequency of visit to each store for each category is a complex periodic vehicle routing problem. Sophisticated optimization models have been built to determine the time of visits for each stores; these optimized visit frequencies are expected to result in a reduction of 10% (equivalent to $150 million) in transportation cost.
- **Pharmacy Department Register Simulation**: Each store has a different volume in terms of patients visit and service time. The simulation model, based on queueing theory, was created to adjust the number of registers to be installed in each store, as a function of service time and customer volume.
- **Call Center Staff Scheduling**: I developed a call center staff scheduling model using genetic algorithm, to minimize under-coverage rate.

Wright State University **Dayton, OH**
Graduate Research and Teaching Assistant *2011–2017*

- **PhD Dissertation: Data mining techniques and mathematical models for the optimal scholarship allocation problem for a state university.** The research uses classification algorithms to find matriculation and graduation rate by varying scholarships. Then the optimization model was developed to optimize revenue under budget, and fairness constraints. This research has prompted the university wide scholarship redesign, the APS calculator, see Website. This project has resulted a 11% (2014) and 13.9% (2015) increase in direct admit students, which translates into a %5 to %10 million dollars of revenue increase for WSU annually.
- **Teaching Assistant:** 1. Intro of Data Mining and Applications. 2. Intro of Operations Research Models.

- **Dissertation Committee:**
  1. Xinhui Zhang(Supervisor): Head of supply chain and operations research at Alibaba. Three times INFORMS Franz Edelman Award finalist.
  2. Pratik Parikh: Chair of Dept of Industrial Engineering, University of Louisville.
  3. Nan Kong: Purdue University.
  4. Caroline Cao: University of South Australia/IMT Atlantique.
  5. Subhashini Ganapathy: Chair of Dept of Biomedical, Industrial & Human Factors Engineering, Wright State University.

Pro Bono

Operations Research/Data Science Consultant

- **Cincinnati Public School Bus Routing and Bell Time Optimization**: I wrote the core optimization algorithm using meta-heuristics, Julia with CBC/Gurobi solver. The result is comparable to the MIT's awarded solution for the Boston public school system.
- **NYC Dog Care Stores Weekly Staff Scheduling Optimization**: I helped a dog care store with 5 locations to create a weekly staff scheduling system with various constraints such as: staff schedule preference, locations preference, demand coverage, and cross-skills satisfaction.