## Contact

www.linkedin.com/in/yuxing-
hou-86032059 (LinkedIn)

## Top Skills

Algorithms

Matlab

R

## Certifications

SAS Certified Advanced
Programmer for SAS 9

SAS Certified Base Programmer for
SAS 9

## Honors-Awards

Graduate Assistantships (2012, 2013
and 2014)

Outstanding Graduate

Best Bachelor's Thesis Award

Second Prize Winner

Merit Student (2009, 2010 and 2011)

# Yuxing Hou

Data Scientist at Ingram Micro
Orange County, California Area

## Summary

Summary:
Solid background in statistics, machine learning and optimization.
Substantial teamwork experience, good communication and
analytical thinking skills.

Specialities:
- Bayesian statistics, Sequential Monte Carlo, Time series analysis.
- Anomaly/novelty detection, supervised and unsupervised learning,
text mining.
- Statistical Modelling: Generalized linear models, Mixed-effects
models, Cox proportional hazards model.
- Optimization: Linear programming, Second-Order Cone (Quadratic)
programming, Integer programming.
- Proficient software skills: R, Python(Scikit-learn, TensorFlow,
Pandas, NLTK, Numpy, XGBoost), Spark, SAS, MATLAB, SQL,
Weka,  AMPL, Mosek.

---

## Experience

**Ingram Micro**
3 years 2 months

Data Scientist
July 2018 - Present (1 year 9 months)

Predictive Analyst
February 2017 - June 2018 (1 year 5 months)
Orange County, California Area

**Center for Computer-Aided Design, University of Iowa**
4 years 10 months

Graduate Research Assistant
August 2012 - May 2017 (4 years 10 months)
Iowa City, Iowa Area

- Built up a new Bayesian framework for process monitoring in short production runs of statistical
process control. Outperformed existing methods in both robustness and computational efficiency.
Experiments were conducted in MATLAB.
- Developed a new sequential Monte Carlo method for online steady-state detection. Outperformed existing online methods in both detection accuracy and computational efficiency. Experiments were conducted in MATLAB.
- Led a group of three to develop a deep learning algorithm for handwritten-digit classification. Programmed in Python (mainly using TensorFlow).
- Worked in a group of four to model the relationship between mid-western real estate sales and various characteristics by using different penalized regression methods (Lasso, Ridge, Elastic Net). Performances were compared via cross-validation. Experiments were conducted in R.
- Led a group of four to develop a new strategy to deal with finite upper bounds for general
linear optimization problems, reducing average time cost by roughly 50% while maintaining the
same optimization performance compared to the built-in MATLAB function. Programmed in
MATLAB.

## Graduate Teaching Assistant
August 2012 - May 2014 (1 year 10 months)
Iowa City, Iowa Area

- Conducted office hours, graded homework and provided feedback on weekly assignments for Operations Research and Simulations with Arena.

## Monsanto Company
Data Scientist Intern
January 2016 - June 2016 (6 months)
Greater St. Louis Area

- Performed sparse clustering (clustering with feature selection) to segment macro-environmental zones for some major crops (corn, soybean, cotton) in the US, with selection of important features affecting the clustering. Algorithm was executed using High Performance Computing (HPC). Results are loaded into Monsanto system for use.
- Applied different machine learning algorithms to investigate the relationship between cotton stress and its influential variables to help improve efficiency of cotton breeding. Algorithm was programmed
using Python and Spark.

Kaggle
Kaggle Competitor: Springleaf Marketing Response
August 2015 - October 2015 (3 months)

- Conducted training process using various classifiers: random forest, gradient boosting trees,
support vector machine and neural network. Performed parameter tuning and probability calibration to improve their prediction accuracy.
- Implemented a nonlinear optimization (Sequential least square quadratic programming) to ensembling different classifiers to achieve better prediction power.
- Programmed in R and Python (Scikit-learn, TensorFlow, Pandas, Numpy, Scipy).
- Final private leaderboard ranking: top 5% among 2225 worldwide teams.

Huazhong University of Science and Technology
Undergraduate Research Assistant
September 2008 - June 2012 (3 years 10 months)
Wuhan, Hubei, China

---

## Education

University of Iowa
Doctor of Philosophy (Ph.D.), Industrial Engineering · (2012 - 2017)

University of Iowa
Master of Science (M.S.), Statistics · (2012 - 2015)

Huazhong University of Science and Technology
Bachelor's degree, Mechanical Engineering · (2008 - 2012)