John Koo

Education

2022 (expected) **Ph.D., Statistical Science**,
Indiana University, Bloomington, IN.
- Advisors: Michael Trosset and Minh Tang
- Dissertation: Latent Structure Block Models and Community Detection

2015 **M.S., Mathematics**,
Tarleton State University, Stephenville, TX.
- Advisors: Bert Little and Jesse Crawford
- Thesis: Statistical Methods for Crop Yield Monitoring via Remote Sensing

2012 **B.A., Mathematics and Physics**,
New York University, New York, NY.

Research Interests

Statistical inference on graphs and networks, computational statistics, dimensionality reduction, statistical pattern recognition, manifold learning.

Publications and Preprints

2021 J. Koo and M. Tang and M. W. Trosset. Popularity Adjusted Block Models are Generalized Random Dot Product Graphs. Submitted for publication. arXiv preprint available at `https://arxiv.org/abs/2109.04010`.

Talks and Presentations

2021 **Connecting the Popularity Adjusted Block Model to the Generalized Random Dot Product Graph**,
Symposium on Data Science and Statistics, Virtual.

2021 **Introduction to Git and GitHub**,
Machine Learning for Research Club at IU, Bloomington, IN.

Software

R package `osc`: An implementation of orthogonal spectral clustering for community detection in linear latent stucture block models. Available at `https://github.com/johneverettkoo/osc`.

Department of Statistics, Indiana University
johnkoo@iu.edu • *johneverettkoo.github.io*
johneverettkoo

Teaching Experience

2018 **Associate Instructor**,
Indiana University, Bloomington, IN.
- Introduction to Business Statistics

2017 - now **Teaching Assistant**,
Indiana University, Bloomington, IN.
- Statistics for the Life Sciences
- Introduction to Statistical Theory
- High Dimensional Data Analysis
- Introduction to Statistical Computing
- Applied Statistical Computing
- Time Series Analysis

Research Experience

2017 - now **Graduate Research Assistant**,
Department of Statistics, Indiana University, Bloomington, IN.
- Proved connections between kernel k-means clustering and normalized and ratio cut algorithms.
- Investigated the effects of data selection for training image recognition models using convolutional neural networks.
- Proved connections between various network models to motivate and develop algorithms for community detection and parameter estimation.

2014 - 2015 **Graduate Research Assistant**,
Department of Mathematics, Tarleton State University, Stephenville, TX.
- Collaboration with the Center for Agribusiness Excellence.
- Developed statistical and machine learning models for crop monitoring and crop insurance fraud detection.

2011 **Undergraduate Research Assistant**,
Courant Institute of Mathematical Sciences, New York University,
New York, NY.
- Investigated steady-state behavior of zero-temperature Ising models via Monte Carlo simulation.

Industry Experience

2020 - now **Data Science Consultant**,
Freelance, Bloomington, IN.
- Developed experimental design procedures for crop field experiments involving variable seeding rates, fertilizer application rates, pesticide application rates, etc.
- Developed algorithms for analyzing real-time monitoring data from planters, sprayers, and combines.
- Built a prototype front-end webapp for customers to design their own crop field experiments.

2019 **Data Science Intern**,
The Climate Corporation, Seattle, WA.
- Trained convolutional neural networks to predict failed crop fields using high resolution spatiotemporal data.

2018 **Data Science Intern**,
The Bee Corp, Bloomington, IN.
- Trained convolutional neural network models to predict beehive health from thermal images.
- Built a front-end webapp for customers to analyze thermal images of beehives.

2015-2017 **Data Analyst**,
The Climate Corporation, San Francisco, CA.
- Developed ETL pipelines for spatiotemporal data of crop field experiments.
- Developed algorithms for analyzing real time monitoring data from planters, sprayers, and combines.
- Developed parameter optimization methods for crop and soil process models by comparing model outputs against data from real time crop and soil sensors.
- Developed and maintained R packages for ingesting, analyzing, and visualizing crop and field data.
- Served as the science team's "R expert".
- Maintained a wiki and repository for crop field trial data. Supported the science team in accessing, analyzing, and interpreting crop field trial data.

2014-2015 **Research Intern**,
Center for Agribusiness Excellence, Stephenville, TX.
- Collaboration with Tarleton State University.
- Developed statistical and machine learning models for crop monitoring and crop insurance fraud detection.

2014 **Software Test Engineer**,
DEKA Research and Development Corp., Manchester, NH.
- Developed and implemented automated software testing procedures for home dialysis machines.

2012-2013 **Mathematics Fellow**,
Match Education, Lawrence, MA.
- Collaboration with Lawrence Public Schools.
- Taught high school level mathematics courses.
- Taught MCAS test prep courses.