Tayler A. Blake

Columbus, OH
☎ *(740) 607-9508*
✉ *tayler.a.blake@gmail.com*
⌨ *http://taylerablake.github.io*

Education

May 2018 **Ph.D. Statistics**, *The Ohio State University*, Columbus, Ohio.
Advisor: Yoonkyung Lee

January 2010 **M.S. Statistics**, *The Ohio State University*, Columbus, Ohio.

May 2007 **B.A. Mathematics, Computer Science**, *Capital University*, Columbus, Ohio.

Experience

Data Scientist, Privacy Engineer

Ketch, *San Francisco, CA*.

September 2021-present

Leads research in the quantification of disclosure risk and serves as engineering's privacy expert. As one of the first members of the company's product and technology organization, lead backend development of Ketch's core data governance product including its native integration with Snowflake, AWS Redshift, Oracle, MySQL, Postgres, and MariaDB database technologies. Owner of the data science efforts supporting Ketch's software for sensitive and personal data discovery and classification.

Senior Research Scientist

Immuta, *Columbus, OH*.

September 2020- September 2021

Senior member of research team driving product feature development for data access and governance software. Lead research and product development of a comprehensive policy-agnostic disclosure risk and utility measurement framework for facilitation of automating privacy policy recommendations for maximum data utility for given disclosure risk tolerance levels.

Lead efforts in designing and prototyping performance testing for the Immuta/Databricks native integration.

Data Scientist

Redjack, *Columbus, OH*.

September 2019- September 2020

Consulting data scientist supporting federal government cyber defense efforts. Leveraged supervised machine learning techniques to identify spam email campaigns, specifically exploratory data analysis techniques for random forests to understand the best features for discriminating between spam and non-spam emails.

Data Scientist

Root Insurance, *Columbus, OH*.

April 2019-
September
2019

Data scientist supporting teams responsible for customer acquisition and development of pricing plans. Designs and analyzes A/B tests and develops conversion and retention models. Such models serve to forecast the impact of changes in pricing methodologies on KPIs which measure the fundamental health of the business, including loss ratio and total bound premium. Integrates with product teams and actuaries at every stage of the data science life cycle, from problem formulation and data collection to model deployment and communication of insights generated from analysis.

Senior Data Scientist

Information Control Company, *Columbus, OH.*

January
2017-March
2019

Consultant serving as a senior data scientist, leading data science team members alongside delivery team members and data engineers. Responsible for planning, designing, and executing data science projects addressing a variety of business questions for clients across several business verticals.

Internally at ICC, contributed to Advanced Analytics development seminars and coordinated the Advanced Analytics journal review, curating and leading discussions about current literature in statistics, data science, and machine learning and its application to ICC client problems.

Machine Learning Specialist

Pillar Technology, *Columbus, OH.*

February
2016-
December
2016

Machine learning specialist on a team responsible for designing and developing adaptive software for an IoT product to be embedded in a line of luxury vehicles. Responsible for setting expectations given the available resources such as data, available in-device storage space, and computational constraints with both internal team and with the client.

Machine learning specialist on a team responsible for designing and developing adaptive software for an IoT product to be embedded in a line of luxury vehicles. Responsible for setting expectations with both internal teams and with the client given the available resources such as data, in-device storage space, and computational constraints. Challenged to design and implement machine learning models in the absence of data a priori, leveraging regularization methods and programmatic model selection technics. Using Agile best practices, owned data science efforts from initial exploratory analysis to production deployment of models in Java and Scala.

Spearheaded components of project planning pertaining to data and the corresponding infrastructure necessary for data collection, storage, and modeling at scale. Lead efforts to establish client trust in data modeling and algorithms, a new operating space for Pillar.

Data Scientist

Store Development, Starbucks Coffee Company, *Seattle, WA.*

May 2014 -
November
2015

Data scientist on a team serving the company in market planning and strategy. Utilized a variety of statistical and machine learning methods, both supervised and unsupervised such as penalized regression and classification, generalized linear models, ensemble methods including boosting, bagging, and random forests, with applications of the latter to both classification and clustering.

Estimated causal impact of a variety of interventions, such as competitor store openings, pricing changes, or new product launches on store performance using Bayesian structural time series models and summarizes comparisons between this methodology and prior approaches, including the traditional difference-in-differences estimators.

Cleaned and analyzed heterogeneous data from several different sources having both potential spatial and temporal components which pose challenges in aggregation and summarization.

Adjunct Statistics Instructor

Department of Mathematics, Columbus State Community College, *Columbus, Ohio.*

August 2013 - January 2014

Lecturer for an introductory statistics course for undergraduate students, Statistics 1350. Non-instruction responsibilities included curriculum and assessment development, including lecture presentations and online learning tools and learning assessments.

Graduate Research Assistant

Comprehensive Cancer Center, The Ohio State University, *Columbus, Ohio.*

October 2011 - August 2012

My responsibilities included analysis of large microarray data sets, in particular utilizing data mining and dimension reduction techniques to find genetic markers in leukemic patients, sharing results and collaborating with medical professionals to both direct further laboratory investigation as well as further statistical investigation.

Graduate Research Assistant

Nationwide Center for Advanced Customer Insights, The Ohio State University, Nationwide Insurance, *Columbus, Ohio.*

June 2010 - June 2011

My responsibilities included work on projects modeling agency behavior using high dimensional demographic and marketing data, specifically by modeling survival times using Cox proportional hazards models with both static and time-varying coefficients. Modeling was done with an emphasis on building parsimonious, interpretable models. I was responsible for presenting results in a corporate setting to high level company executives with motive to encourage and motivate business decisions and action.

Graduate Teaching Assistant

Department of Statistics, The Ohio State University, *Columbus, Ohio.*

August 2007 - May 2014

Lecturer for a statistics course for undergraduate engineering students, Statistics 427.

Lecturer for an introductory statistics course, Statistics 145 and a statistics courses for undergraduate engineering students, Statistics 427 and 3470.

Lab instructor for an introductory statistics course, Statistics 145, 245, 133, and 2480.

Research Interests

My research interests include nonparametric function estimation, particularly utilizing reproducing kernel Hilbert space methods; my dissertation work focuses on nonparametric estimation of large covariance matrices through the estimation of covariance functions, specifically in the case of irregularly spaced and sparsely sampled data. We propose an estimation framework with unconstrained optimization through a specific decomposition of a covariance matrix. This parameterization renders covariance estimation as the estimation of a varying coefficient regression model, which allows us access to machinery typically used in the classic function estimation setting for estimating a positive definite covariance matrix.

Computing Skills

Statistical/Analysis Software: R, RMarkdown, R Shiny, SparkR, basic knowledge of Apache Spark

Programming Languages: MySQL, PostgreSQL, basic knowledge of Python and HTML

Proficient in Git and version control, MS Office, LaTeX, project management software including Jira and Confluence

Presentations

August 2018	Joint Statistical Meetings	*Smoothing spline ANOVA models for nonparametric covariance estimation for longitudinal data*
August 2017	Joint Statistical Meetings	*Nonparametric covariance estimation for longitudinal data via tensor product smoothing*
October 2016	Wards Auto User Experience Conference	*The Machines Are Coming: Will Algorithms Replace Designers in the UX World?*
August 2012	Joint Statistical Meetings	*Nonparametric Covariance Estimation for Functional Data with Shrinkage Toward Stationary Models*

Honors and Awards

June 2010	Ohio State University Department of Statistics Teaching Assistant of the Year Nominee
June 2009	Ohio State University Department of Statistics Teaching Assistant of the Year
June 2008	Ohio State University Department of Statistics Teaching Assistant of the Year Nominee

Service

2007 - 2014	**Member, Graduate Recruitment Committee**, *Department of Statistics, The Ohio State University.*
2007 - 2012	**Chair, Graduate Recruitment Committee**, *Department of Statistics, The Ohio State University.*

Hobbies and Interests

Competitive distance running, cooking, baking, reading, practicing yoga

References

Available upon Request