

Research Statement

Karsten T. Maurer

At my academic core I am comprised of three intertwined identities. I am a data scientist. I am a statistics educator. I am an inter-disciplinary thinker. The foundations of my academic-self are built atop my liberal undergraduate education at Minnesota Morris. There I was instilled with a driving curiosity in many subjects and the realization that data skills were universally valuable across disciplines. My graduate degrees from Iowa State pushed and challenged me to deeply learn statistical theory and methodology. In truth, I stuck with the pursuit of my PhD for three reasons: (1) I am stubbornly persistent, (2) I was given the chance to teach, I loved it and wanted the degree so that I could be a professor, and (3) the data science skills I was learning along the way were seriously cool. As an assistant professor of statistics at Miami University, my passion for higher education, data science and inter-disciplinary application of statistics have been galvanized. I have had the opportunity to indulge my wide-ranging curiosity through research collaborations with colleagues all over campus. I also had the opportunity to use my data science background to extend into machine learning research in a collaborative partnership with Dr. Walter Bennette at the Air Force Research Lab (AFRL). I risk the drudgery of presenting my academic biography to begin my research statement because it reveals *why* my research is focused on work in teaching methodology for higher education, machine learning methodology, and statistical support for inter-disciplinary projects.

My research in teaching methods for higher education has included evaluations of both discipline-based and general educational practices. In statistics education, I have conducted comparative assessment into the efficacy of several curricular components on achieving learning outcomes and student engagement. These include comparisons of traditional vs. simulation-based inference curricula (Maurer and Lock, 2016), instructor vs. publisher created homework structures (Swart et al., *submitted 2019*), and a tactile dice game vs. word problem lab on independent random variables (Maurer et al., *submitted 2019*). I have also lead a group of statistics educators at Miami in a curricular content audit of our coverage of p-values in light of the ASA p-value statement (Maurer et al., 2019). I am also interested in the development of discipline-specific digital teaching tools, such as my *Shiny Database Sampler* app that allows introductory students to practice sampling techniques on large databases (ex. Census Data) as their population, then plot or export the sample data for further use (Maurer and Hofmann, *submitted 2019*). I have also been involved in general cross-disciplinary research in higher education through projects to evaluate and document the change in student performance after implementing digital proctoring software for online

exams; spoiler - it is *very* important to use proctoring software for online exams (Alessio and Maurer 2018; Alessio et al. 2018; Alessio et al. 2017). Along with innovative thought into new pedagogical approaches, I believe that formal comparative analysis is necessary to establish evidence-based best practices. I will continue to analytically evaluate the efficacy of innovative efforts I make in my classrooms and pursue publication to share the successful elements with the educational community. I will also continue to expand my engagement in professional exchange with statistics and data science educators from many institutions in order to benefit from the cross-pollination of pedagogical development.

My machine learning research has primarily centered on diagnostic methods for classification models. I work on developing active learning algorithms for efficiently searching large sets of unlabeled instances for cases that highlight deficiencies in a classifier, which may occur from training set bias or from domain shift. Previous research in this area has focused on identifying “unknown unknowns” – highly confident but misclassified instances – using utility function optimization to suggest the most valuable cases on which to spend a limited resource budget when adding labels. There is a fundamental oversight in these methods; they assume that all high-confidence errors demonstrate a deficiency with the classifier. Whereas in the work with my collaborator, Dr. Bennette, our facility locations utility function is adapted to account for the *expected* number of misclassifications that *should* be occurring at a rate described by the confidence values, thus providing a labeling algorithm that seeks to identify the unlabeled instances on which the classifier is most over-confident (Maurer and Bennette, 2019). Our facility locations search algorithm updates actively as we add labels to the selected instances, and while it tends to improve with each new case, the improvement is slow in initial sampling. We are following up with explorations of an algorithm that can be used to kick-start these searches by identifying the “low-hanging fruit” – instances where the classifier has high confidence, but the feature values suggest that they are tenuously classified. We do this through adversarial learning techniques that identify the minimal change to the feature values that would result in a change in the predicted class – a minimal adversarial distance. We note that these adversarial distances and predicted class confidence values can be generated without human supervision, and thus we can algorithmically identify points that have low adversarial distances relative to their high predicted confidence values in a large unlabeled set. In our current research, we find that when we used such points to initialize our query, we find over-confident cases much more efficiently than while using previous methods for searching for unknown unknowns. We have worked with this adversarial distance search method for cases involving image and text classifiers, and are currently working to generalize the method for all types of classifiers. I anticipate that this work will evolve in the future, but in the near-term will continue in the areas of

active learning, classifier diagnostics and confidence calibration.

Having the local reputation as an applied statistician with data science skills and an eclectic set of interests has lead to many opportunities for collaborative, inter-disciplinary research. In my previous projects, I have worked collaboratively with researchers in the fields of actuarial science (Miljkovic et al., 2018), machine learning (Maurer and Bennette, 2019), economics (Anderson et al., 2013), education (Alessio and Maurer, 2018), and kinesiology (Alessio et al. 2018; Alessio et al. 2017). I am also currently acting as the statistical collaborator on the following research projects: a meta analysis of beaver dam ecosystem impacts with a team of biologists, ecologists and EPA specialists; an analysis of lab-tests on baby food consistency across manufacturers with speech pathologists and a mechanical engineer; a permutation-based testing method for comparing eye-tracking patterns in A-B testing of webpage designs with a marketing researcher; and a web app for interactive data visualization of the National Lakes Assessment data with biologists. Interacting with the wealth of experience and knowledge of my collaborators in this inter-disciplinary research has helped me grow my abilities in research, data analysis and communication. I will continue to engage with collaborators that have interesting research problems and are seeking the data abilities to answer them.

A substantial part of my research program is conducted with contributions from my undergraduate and graduate students. I believe that getting students involved with scholarly research presents them with the opportunity to learn topics more deeply, and practice problem solving more intensely, than they will experience in most classes. I am however intentional about the projects and students that I choose to advise, because I respect the time and effort that are required for the research mentorship to be valuable for both the advisor and student. I have worked on undergraduate research projects with six students. Yuexi Wang assisted me in designing and running computational experiments for comparing bin-weighted ensemble classifiers. I co-advised Robert Garrett and Austin Nar as they developed the `ggvoronoi` package on CRAN, and published an associated article for the Journal of Open Source Software (Garrett et al., 2018). Robert, Dr. Tom Fisher and myself have also started the groundwork on a paper about general applications of voronoi diagrams in data science. I co-advised Alison Tuiyott, Ben Schweitzer, Robert Garrett and Lydia Carter as they worked on two unique projects for the ASA Data Expo Competition, and have been continuing to advise Ben and Alison on drafting their papers to Computational Statistics, which were invited due to their 2nd and 3rd place finishes at the Data Expo, respectively. While all of the master's students at Miami are required to complete a capstone project that demonstrates their abilities, none are required to work toward publication. I have advised and co-advised nine statistics students and one computer science student on their masters

capstones. Of these, four have contributed substantially toward papers that either have been submitted or are in preparation. David Swart conducted educational research on homework design for introductory statistics. David has since been hired as an instructor in our department and has submitted the manuscript of his work to the Statistics Education Research Journal, coauthored by Lynette Hudiburgh and myself. Bunyod Tusmatov and Sally Dufek have both worked toward generalizing adversarial distance search methods with myself and Dr. Bennette. Josh Morgan is working toward a manuscript, along with his co-advisor Dr. John Femiani and myself, on using an active learning framework and customized loss functions to improve β -Variational auto-encoder image classifiers, and has submitted initial results to the student poster section of the AAAI conference. I am extremely proud of the accomplishments of my students. I will continue to welcome a diverse group of ambitious, talented and capable students into my projects, and help them as they develop their own projects. Students greatly enhance both the quality and joy I find in doing research.

The research I conduct enhances me as a teacher, as a data professional and as a thinker. I am enthusiastic to continue my educational, machine-learning and inter-disciplinary work. I grow the most as a scholar when working and learning with others, therefore much of this work will continue to be done collaboratively with students, colleagues and professional partners.

References

- Alessio, H., Malay, N., Maurer, K., Bailer, J., and Rubin, B. (2018), “International Review of Research in Open and Distributed Learning,” *Interaction*, 19.
- Alessio, H. and Maurer, K. (2018), “The Impact of Video Proctoring in Online Courses.” *Journal on Excellence in College Teaching*, 29, 183–192.
- Alessio, H. M., Malay, N., Maurer, K., Bailer, A. J., and Rubin, B. (2017), “Examining the effect of proctoring on online test scores.” *Online Learning*, 21, 146–161.
- Anderson, J., Burks, S. V., Carpenter, J., Götte, L., Maurer, K., Nosenzo, D., Potter, R., Rocha, K., and Rustichini, A. (2013), “Self-selection and variations in the laboratory measurement of other-regarding preferences across subject pools: evidence from one college student and two adult samples,” *Experimental Economics*, 16, 170–189.
- Garrett, R. C., Nar, A., Fisher, T. J., and Maurer, K. (2018), “ggvoronoi: Voronoi Diagrams and Heatmaps with ggplot2.” *J. Open Source Software*, 3, 1096.
- Maurer, K. and Bennette, W. (2019), “Facility Locations Utility for Uncovering Classifier Overconfidence,” *arXiv preprint arXiv:1810.05571*.
- Maurer, K., Hudiburgh, L., Werwinski, L., and Bailer, J. (2019), “Content Audit for p-value Principles in Introductory Statistics,” *The American Statistician*, 73, 385–391.
- Maurer, K. and Lock, D. (2016), “Comparison of learning outcomes for simulation-based and traditional inference curricula in a designed educational experiment,” *Technology Innovations in Statistics Education*, 9.
- Miljkovic, T., Miljkovic, D., and Maurer, K. (2018), “Examining the impact on mortality arising from climate change: important findings for the insurance industry,” *European Actuarial Journal*, 8, 363–381.