

Alexej Gossmann

Bioinnovation PhD Program
Tulane University

Web: alexejgossmann.com
Email: agossman@tulane.edu

AREAS OF INTEREST

Development of statistical and machine learning techniques for feature selection and prediction on big high-dimensional datasets, with focus on understanding how false positive findings can arise, and how they can be avoided. Applications in genomics and neuroimaging. Translation of the statistical and machine learning methodology into a useful (software) product.

EDUCATION

PhD, Bioinnovation
Tulane University, New Orleans, Louisiana, in progress

Doctoral research, Mathematics
Tulane University, New Orleans, Louisiana, through January 2017
GPA: 3.975

MS, Statistics
Tulane University, New Orleans, Louisiana, May 2014
GPA: 3.975

BS, Mathematics
Technische Universität Darmstadt, Darmstadt, Germany, May 2012
GPA: 4.0

EXPERIENCE

- Research assistantship at The Multiscale Bioimaging and Bioinformatics Laboratory at Tulane University under the supervision of Dr. Yu-Ping Wang. Spring 2015 – present.
- Student Intern doing machine learning research in the Division of Imaging, Diagnostics, and Software Reliability (CDRH/OSEL/DIDSR) at the U.S. Food & Drug Administration (FDA). May – August 2017.
- Instructor, co-teacher, or teaching assistant for various undergraduate statistics, calculus, and real analysis courses at Tulane University and Technische Universität Darmstadt, Fall 2010 - Fall 2014.
- Student developer for Google Summer of Code 2015. Project: Adding Linear Mixed Effects Models Support to SciRuby. Supervised by the Ruby Science Foundation. May – August 2015.
- Mentor for Google Summer of Code 2016. Project: Categorical data support for Daru, Statsample and Statsample-glm. May – August 2016.
- Mentor for Google Summer of Code 2017. Project: Creating the fastest math libraries for Ruby by using the GPU through OpenCL and ArrayFire. May – August 2017.
- Research assistantship under the supervision of Dr. Oleksandr Gromenko at the Department of Mathematics, Tulane University, working on novel statistical methods for the analysis spatio-temporal processes.. June – August 2014.

- President of the SIAM student chapter at Tulane University. September 2014 – September 2016.
- Main organizer of the Graduate Student Colloquium at the Department of Mathematics, Tulane University. September 2014 – September 2016.

SKILLS

Computer skills: R, Ruby, C++, Python, Matlab, L^AT_EX, Linux, git and github, HTML, CSS.

Language Knowledge: Bilingual in German and Russian, fluent in English, basic knowledge of French.

HONORS & AWARDS

- Reviewers Choice Abstract (top 10% of poster abstracts), 66th Annual Meeting of The American Society of Human Genetics, October 19, 2016, Vancouver, Canada.
- SIAM Student Chapter Certificate of Recognition, SIAM (Society for Industrial and Applied Mathematics), April 2015.

PUBLICATIONS

- [1] Alexej Gossmann, Aria Pezeshk, and Berkman Sahiner. Test data reuse for evaluation of adaptive machine learning algorithms: Over-fitting to a fixed “test” dataset and a potential solution. In *Proceedings of SPIE: Medical Imaging 2018*, 2018.
- [2] Alexej Gossmann, Pascal Zille, Vince Calhoun, and Yu-Ping Wang. Fdr-corrected sparse canonical correlation analysis with applications to imaging genomics. *arXiv preprint arXiv:1705.04312*, 2017 **under review**.
- [3] Alexej Gossmann, Shaolong Cao, Damian Brzyski, Lan-Juan Zhao, Hong-Wen Deng, and Yu-Ping Wang. A sparse regression method for group-wise feature selection with false discovery rate control. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2016 **under review**.
- [4] Damian Brzyski, Alexej Gossmann, Weijie Su, and Małgorzata Bogdan. Group slope-adaptive selection of groups of predictors. *arXiv preprint arXiv:1610.04960*, 2016 **under review**.
- [5] Shaolong Cao, Huaizhen Qin, Alexej Gossmann, Hong-Wen Deng, and Yu-Ping Wang. Unified tests for fine scale mapping and identifying sparse high-dimensional sequence associations. *Bioinformatics*, 2015.
- [6] Mimi C Sammarco, Jennifer Simkin, Alexander J Cammack, Danielle Fassler, Alexej Gossmann, Luis Marrero, Michelle Lacey, Keith Van Meter, and Ken Muneoka. Hyperbaric oxygen promotes proximal bone regeneration and organized collagen composition during digit regeneration. *PLoS one*, 10(10), 2015.
- [7] Shaolong Cao, Huaizhen Qin, Alexej Gossmann, Hong-Wen Deng, and Yu-Ping Wang. Unified tests for fine scale mapping and identifying sparse high-dimensional sequence associations. In *Proceedings of the 6th ACM Conference on Bioinformatics, Computational Biology and Health Informatics, BCB '15*, pages 241–249, New York, NY, USA, 2015. ACM.
- [8] Alexej Gossmann, Shaolong Cao, and Yu-Ping Wang. Identification of significant genetic variants via slope, and its extension to group slope. In *Proceedings of the 6th ACM Conference on Bioinformatics, Computational Biology and Health Informatics, BCB '15*, pages 232–240, New York, NY, USA, 2015. ACM.

CONFERENCE PRESENTATIONS

- SlopeCCA and gslopeCCA: sorted L1 penalized canonical correlation analysis (Abstract/Program #2803W). Presented at the *66th Annual Meeting of The American Society of Human Genetics*, October 19, 2016, Vancouver, Canada.
- Identification of Significant Genetic Variants via SLOPE, and its Extension to Group SLOPE (Abstract/Program #1343F). Presented at the *65th Annual Meeting of The American Society of Human Genetics*, October 9, 2015, Baltimore, MD.
- Identification of Significant Genetic Variants via SLOPE, and its Extension to Group SLOPE. The *6th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics*, Atlanta, GA, September 2015.

SOFTWARE

- `grpSLOPE` – Group SLOPE (Group Sorted L1 Penalized Estimation) is a penalized linear regression method that is used for adaptive selection of groups of significant predictors in a high-dimensional linear model. The Group SLOPE method can control the (group) false discovery rate at a user-specified level. Project repository: <https://github.com/agisga/grpSLOPE>. CRAN: <https://cran.r-project.org/web/packages/grpSLOPE/index.html>.
- `grpSLOPEMC` – This is an extension package to the R package `grpSLOPE`. It contains Monte Carlo based methods for the estimation of the regularizing sequence. Project repository: <https://github.com/agisga/grpSLOPEMC>.
- `mixed_models` – Fit statistical linear models with fixed and random effects in Ruby. Project repository: https://github.com/agisga/mixed_models. RubyGems: https://rubygems.org/gems/mixed_models.
- In my free time I like to contribute to open source software projects. Visit my Github page for the projects that I contribute to: <https://github.com/agisga>.