

# SEBASTIAN RASCHKA

## Curriculum Vitae

### Contact Information

---

#### Private

**Address** [REDACTED]  
**Phone** [REDACTED]  
**Email** [REDACTED]  
**Website** <https://sebastianraschka.com>

#### Work

**Office Address** [REDACTED]  
**Office Phone** [REDACTED]  
**Work Email** [sraschka@wisc.edu](mailto:sraschka@wisc.edu)  
**Website** <http://pages.stat.wisc.edu/~sraschka>

### Professional Experience and Employment History

---

**Aug 2018-present** University of Wisconsin-Madison  
*Assistant Professor of Statistics (tenure-track)*

Machine learning and deep learning research with focus on ordinal regression and applications to computational biology. Teaching machine learning and deep learning courses in the Department of Statistics. Tenure-track 2018-2024.

- Feb 2018-  
July 2018** Protein Structural Analysis & Design Lab Prof. Kuhn  
*Temporary Research Consultant*
- Developed predictive models and *in silico* methods to aid the discovery of potential inhibitors of cancer metastasis.
- 2012-2017** Protein Structural Analysis & Design Lab Prof. Kuhn  
*Graduate Research Assistant*
- Conducted statistical analyses and developed computational tools that lead to the discovery of clear and strong patterns of chemical group matching preferences for intermolecular hydrogen bonding. The results from these statistical analyses lead to new insights that have wide implications for protein and ligand design.
- Collaborated with experimental biologists to design machine learning models in time-sensitive projects that led to the discovery of receptor antagonists of sea lamprey mating pheromones.
- Developed a hypothesis-driven methodology and Python software framework, Screenlamp, for large-scale virtual screening allowing researchers to identify target molecules of protein receptors at an unprecedented scale in databases of more than 15 million compounds.
- Developed algorithms and a Python/C software package for predicting 3-dimensional protein-ligand binding poses on the time scale of seconds and high predictive accuracy based on graph theory.
- Developed a novel approach, 3D epitope-based virtual screening, for identifying small molecule inhibitors of protein-protein interactions. This new method was applied to screen 10 million commercially available molecules with drug-like properties and identified two candidates that were shown to be successful as inhibitors of cancer metastasis in experimental assays performed by collaborators.
- Managed IT systems in the Kuhn lab, including hardware, software, servers, computing clusters, and Linux workstations.
- May 2012-  
Aug 2012** Computational Pharmaceutical Chemistry Lab Dr. Holger Gohlke  
*Research Internship*
- Research project: *Analysis of protein thermostability using a constraint network approach based on graph theory.*
- Conducted molecular dynamics simulations and analyzed the resulting trajectories to predict the effect of binding-site mutations on the thermostability of various proteins.
- May 2010-  
Oct 2010** Immulab Duesseldorf  
*Internship*

Performed computational and statistical analyses of real-time qPCR data from human saliva samples to diagnose paradontose causing bacteria.

## Education

---

**2012-2017** Quantitative Biology and Biochemistry and Molecular Biology  
Ph.D. - Michigan State University

Thesis title: Uncovering Hidden Patterns of Molecular Recognition

Main research topics:

1) *Use of statistical data mining approaches and development of new software protocols to uncover novel patterns of preferential interactions between proteins and their small molecule binding partners (the Protein Recognition Index and Screenlamp data mining toolkits), for use in enhancing the design of proteins and small molecules for protein engineering and drug discovery.*

2) *Development and application of a new method and software application, Screenlamp, for hypothesis-driven, high-throughput virtual screening of millions of small molecules to discover molecules that inhibit proteins. This has led to our pioneering success in aquatic invasive species control and also has great potential for drug discovery (<https://psa-lab.github.io/screenlamp>).*

3) *Development of machine learning protocols for identifying the chemical groups within a molecule that lead to its biological activity, which is a key step to improving the potency of small molecules as biological inhibitors or activators (<https://github.com/psa-lab/predicting-activity-by-machine-learning>).*

4) *Development of a novel algorithm, SiteInterlock, for predicting how small molecule partners bind to proteins, which showed that rigidity is a hallmark of correctly oriented protein-small molecule complexes. This graph-theoretic analysis is being distributed to the public in well-documented software through GitHub (<https://psa-lab.github.io/siteinterlockl>).*

5) *Invention of 3D epitope-based virtual screening, a novel computational approach for inhibiting protein-protein interactions that enabled the discovery of promising compounds for limiting intestinal cancer metastasis.*

PDF: [https://sebastianraschka.com/pdf/articles-and-preprints/raschka\\_phd\\_thesis.pdf](https://sebastianraschka.com/pdf/articles-and-preprints/raschka_phd_thesis.pdf)

Advisor: Prof. Dr. Leslie A. Kuhn

**2008-2012** B.Sc. Biology Plus International - Heinrich-Heine-Universitaet Duesseldorf

Thesis: *Investigation of potential CLAVATA2 interactors by Fluorescence Resonance Energy Transfer Analysis*

Advisor: Prof. Dr. Ruediger Simon

**2011-2012** Study Abroad - Michigan State University

Research during the 1st semester: *Acetylation-dependent SUMOylation of p53*

Advisor: Prof. Dr. Min-Hao Kuo

Research during the 2nd semester: *Phylogenetic studies of Rhagoletis flies and associated parasitoid Coptera wasps*

Advisor: Prof. Dr. Jim Smith

**1996-2007** High School Diploma - Gymnasium Rheinkamp Europaschule Moers

## Patents

---

- M.D. Basson, L.A. Kuhn, S. Raschka, "Inhibiting FAK-AKT Interaction to Inhibit Metastasis," United States Patent. Appl. No.: PCT/US2018/042919; Pub. No.: WO/2019/018666.

## Top 5 Publications

---

- Sebastian Raschka and Benjamin Kaufman (2020)  
*Machine learning and AI-based approaches for bioactive ligand discovery and GPCR-ligand recognition*. Elsevier Methods, 180, 89–110
- Sebastian Raschka, Alex Wolf, and Leslie A. Kuhn (2018)  
*Protein-ligand interfaces are polarized: Discovery of a strong trend for intermolecular hydrogen bonds to favor donors on the protein side with implications for predicting and designing ligand complexes*. Journal of Computer-Aided Molecular Design, 32(4), 511-528.
- Sebastian Raschka, Nan Liu, Santosh Gunturu, Anne M. Scott, Mar Huertas, Weiming Li, and Leslie A. Kuhn (2018)  
*Facilitating the hypothesis-driven prioritization of small molecules in large databases: Screenlamp and its application to GPCR inhibitor discovery*.
- Sebastian Raschka, Joseph Bemister-Buffington, and Leslie A. Kuhn (2016)  
*Detecting the native ligand orientation by interfacial rigidity: SiteInterlock*. Proteins: Structure, Function and Bioinformatics 84.12: 1888-1901.
- Wenzhi Cao, Vahid Mirjalili, and Sebastian Raschka (2020)  
*Rank Consistent Ordinal Regression for Neural Networks with Application to Age Estimation*. Pattern Recognition Letters. 140, 325-331

## Research Articles

---

- Zhongjie Yu and Sebastian Raschka (2021)  
*Simple Post-Hoc Work Can Improve Supervised and Semi-Supervised Few-Shot Learning.* (under review)
- BD Lee, A Gitter, CS Greene, S Raschka, F Maguire, AJ Titus, MD Kessler, AJ Lee, MG Chevrette, PA Stewart, T Britto-Borges, EM Cofer, K-H Yu, JJ Carmona, EJ Fertig, AA Kalinin, B Signal, TJ Triche, Jr., SM Boca (2021)  
*Ten Quick Tips for Deep Learning in Biology.* (under review)
- Sebastian Raschka (2021)  
*Deeper Learning By Doing: Integrating Hands-On Research Projects Into A Machine Learning Course* (under review)
- Kaiping Chen, Sang Jung Kim, Sebastian Raschka, and Qiantong Gao (2021)  
*Visual Framing of Science Conspiracy Videos: Integrating Machine Learning with Communication Theories to Study the Use of Color and Brightness.* (under review)
- Zhongjie Yu, Lin Liu, and Sebastian Raschka (2021)  
*Few-Shot Learning for Video Object Detection.* (under review)
- Runxin Gao, Kaiping Chen, Sebastian Raschka (2021)  
*Using Language Models For Supervised Learning When Labeled Data Is Scarce: A Case Study Of Classifying Political Communication Texts.* (under review)
- Sebastian Raschka (2021)  
*Deeper Learning By Doing: Integrating Hands-On Research Projects Into A Machine Learning Course.* (accepted at Teaching ML workshop, ECML 2021)
- Wenzhi Cao, Vahid Mirjalili, and Sebastian Raschka (2020)  
*Rank Consistent Ordinal Regression for Neural Networks with Application to Age Estimation.* Pattern Recognition Letters. 140, 325-331
- Zhongjie Yu and Sebastian Raschka (2020)  
*Looking Back to Lower-level Information in Few-Shot Learning.* Information 2020, 11, 7.
- Sebastian Raschka, Joshua Patterson, and Corey J. Nolet (2020)  
*Machine Learning in Python: Main Developments and Technology Trends in Data Science, Machine Learning, and Artificial Intelligence.* Information 2020, 11, 4
- Vahid Mirjalili, Sebastian Raschka, and Arun Ross (2020)  
*PrivacyNet: Semi-Adversarial Networks for Multi-attribute Face Privacy* IEEE Transactions in Image Processing. Vol. 29, pp. 9400-9412, 2020.
- Sebastian Raschka and Benjamin Kaufman (2020)  
*Machine learning and AI-based approaches for bioactive ligand discovery and GPCR-ligand recognition.* Elsevier Methods, 180, 89–110

- Joseph Bemister-Buffington, Alex J. Wolf, Sebastian Raschka, and Leslie A. Kuhn (2020)  
*Machine Learning to Identify Flexibility Signatures of Class A GPCR Inhibition*. *Biomolecules* 2020, 10, 454.
- Vahid Mirjalili, Sebastian Raschka, and Arun Ross (2019)  
*FlowSAN: Privacy-enhancing Semi-Adversarial Networks to Confound Arbitrary Face-based Gender Classifiers*. *IEEE Access* 2019, 10.1109/ACCESS.2019.2924619
- Sebastian Raschka (2019)  
*Automated discovery of GPCR bioactive ligands*. *Current Opinion in Structural Biology* 2019, 55:17–24
- Vahid Mirjalili, Sebastian Raschka, and Arun Ross (2018)  
*Gender Privacy: An Ensemble of Semi Adversarial Networks for Confounding Arbitrary Gender Classifiers*. Proc. of 12th IEEE International Conference on Biometrics: Theory, applications and systems (BTAS 2018), Los Angeles.
- Sebastian Raschka, Shyam K. More, Dinesh Devadoss, Bixi Zeng, Leslie Kuhn, Marc D. Basson (2018).  
*Identification of potential small-molecule protein-protein inhibitors of cancer metastasis by 3D epitope-based computational screening*. *Journal of Physiology and Pharmacology* 69, 2.  
(Journal link: [http://www.jpp.krakow.pl/journal/archive/04\\_18/pdf/10.26402/jpp.2018.2.11.pdf](http://www.jpp.krakow.pl/journal/archive/04_18/pdf/10.26402/jpp.2018.2.11.pdf))
- Sebastian Raschka (2018)  
*MLxtend: Providing machine learning and data science utilities and extensions to Python's scientific computing stack*. *The Journal of Open Source Software* 3.24.  
(Journal link: <http://joss.theoj.org/papers/10.21105/joss.00638>)
- Vahid Mirjalili, Sebastian Raschka, Anoop Namboodiri, and Arun Ross (2018)  
*Semi-adversarial networks: Convolutional autoencoders for imparting privacy to face images*. Proc. of 11th IAPR International Conference on Biometrics (ICB 2018), Gold Coast, Australia. (Best Paper Award)  
(Journal link: <https://ieeexplore.ieee.org/document/8411207/>;  
arXiv preprint: <https://arxiv.org/abs/1712.00321>)
- Sebastian Raschka, Alex Wolf, and Leslie A. Kuhn (2018)  
*Protein-ligand interfaces are polarized: Discovery of a strong trend for intermolecular hydrogen bonds to favor donors on the protein side with implications for predicting and designing ligand complexes*. *Journal of Computer-Aided Molecular Design*, 32(4), 511-528.  
(Journal link: <https://link.springer.com/article/10.1007/s10822-018-0105-2>;  
bioRxiv preprint: <https://www.biorxiv.org/content/early/2018/02/05/260612>)
- Sebastian Raschka, Nan Liu, Santosh Gunturu, Anne M. Scott, Mar Huertas, Weiming Li, and Leslie A. Kuhn (2018)  
*Facilitating the hypothesis-driven prioritization of small molecules in large databases: Screenlamp and its application to GPCR inhibitor discovery*. *Journal of Computer-Aided Molecular Design*, 32(3), 415-433.  
(Journal link: <https://link.springer.com/article/10.1007/s10822-018-0100-7>;  
bioRxiv preprint: <https://www.biorxiv.org/content/early/2018/01/17/249151>)

- Sebastian Raschka (2017)  
*BioPandas: Working with molecular structures in pandas DataFrames*. The Journal of Open Source Software 2.14.  
(Journal link: <http://joss.theoj.org/papers/10.21105/joss.00279>)
- Sebastian Raschka, Joseph Bemister-Buffington, and Leslie A. Kuhn (2016)  
*Detecting the native ligand orientation by interfacial rigidity: SiteInterlock*. Proteins: Structure, Function and Bioinformatics 84.12: 1888-1901.  
(Journal link: <http://onlinelibrary.wiley.com/doi/10.1002/prot.25172/summary>)

## Technical Reports

- Sebastian Raschka (2018). *Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning*. <https://sebastianraschka.com/pdf/manuscripts/model-eval.pdf>
- Sebastian Raschka (2015). *MusicMood: Predicting the mood of music from song lyrics using machine learning*. arXiv:1611.00138.
- Sebastian Raschka (2014). *Naive Bayes and Text Classification I-Introduction and Theory*. arXiv:1410.5329.

## Books

---

- Sebastian Raschka, Vahid Mirjalili. *Python Machine Learning. 3rd Edition*. Birmingham, UK: Packt Publishing, 2019. ISBN: 978-1789955750.
- Sebastian Raschka, Vahid Mirjalili. *Python Machine Learning. 2nd Edition*. Birmingham, UK: Packt Publishing, 2017. ISBN: 978-3958457331. (More than 18,700 copies sold as of July 1st, 2018)
- Sebastian Raschka, David Julian, and John Hearty. *Python: Deeper Insights into Machine Learning*. Birmingham, UK: Packt Publishing, 2016. ASIN: B01LD8K994.
- Sebastian Raschka. *Python Machine Learning*. Birmingham, UK: Packt Publishing, 2015. ISBN: 978-1783555130. (Amazon bestseller in Data Mining 2015 & 2016, Packt bestselling title 2015 & 2016, more than 42,500 copies sold as of July 1st, 2018; ACM Computing Reviews' Notable Computing Books and Articles of 2016)
- Sebastian Raschka. *Heat Maps in R: How-To*. Birmingham, UK: Packt Publishing, 2013. ISBN: 978-1782165644. (more than 500 copies sold)

## Books, Translations into Other Languages

- Sebastian Raschka, Vahid Mirjalili. *Machine Learning mit Python und Keras, TensorFlow 2 und Scikit-learn: Das umfassende Praxis-Handbuch fuer Data Science, Deep Learning und Predictive Analytics (German Transl.)*. Germany: mitp, 2021. ISBN: 978-3747502136.

- Sebastian Raschka, Vahid Mirjalili. *Python masinsko učenje (Bosnian Transl.)*. Bosnia: Kompjuter biblioteka, 2020. ISBN: 978-8673105499.
- Sebastian Raschka, Vahid Mirjalili. *Machine Learning con Python – nuova edizione (Italian Transl.)*. Italy: Apogeo, 2020. ISBN: 978-8850335244.
- Sebastian Raschka, Vahid Mirjalili. *Python Machine Learning. 2nd Edition (Polish Transl.)*. Poland: Helion. ISBN: 978-8328351219.
- Sebastian Raschka, Vahid Mirjalili. *Python Machine Learning. 2nd Edition (Japanese Transl.)*. Japan: Impress Top Gear. ISBN: 978-4295003373.
- Sebastian Raschka, Vahid Mirjalili. *Python Machine Learning. 2nd Edition (German Transl.)*. Frechen, Germany: mitp Verlag, 2017. ISBN: 978-3958457331.
- Sebastian Raschka, Vahid Mirjalili. *Python Machine Learning. 2nd Edition (Spanish Transl.)*. Barcelona, Spain: mitp Verlag, 2019. ISBN: 978-8426727206.
- Sebastian Raschka, Vahid Mirjalili. *Python Machine Learning. 2nd Edition (Korean Transl.)*. South Korea, Gyeonggi: Gilbut Publishing, 2019. ISBN: 978-8426727206.
- Sebastian Raschka. *Python Machine Learning (Russian Transl.)*. Russia: DMK Press, 2017. ISBN: 978-5970604090.
- Sebastian Raschka. *Python Machine Learning (Polish Transl.)*. Poland: Helios, 2017. ISBN: 978-8328336131.
- Sebastian Raschka. *Python Machine Learning (Korean Transl.)*. Korea: Ji & Son, 2017. ISBN: 979-1187497035.
- Sebastian Raschka. *Machine Learning mit Python (German Transl.)*. Frechen, Germany: mitp Verlag, 2016. ISBN: 978-3958454224.
- Sebastian Raschka. *Machine learning con Python. Costruire algoritmi per generare conoscenza (Italian Transl.)*. Italy: Apogeo, 2016. ISBN: 978-8850333974.
- Sebastian Raschka. *Python Machine Learning (Japanese Transl.)*. Japan: Impress Top Gear, 2016. ISBN: 978-4844380603.
- Sebastian Raschka. *Python Machine Learning (Mainland Chinese Transl.)*. Chinfong, China: Mechanical Engineering Press, Ltd., 2017. ISBN: 978-7111558804.
- Sebastian Raschka. *Python Machine Learning (Traditional Chinese Transl.)*. New Taipei City, Taiwan: DrMaster Press Co., Ltd., 2016. ISBN: 978-9864341405.

## Book Chapters

- Sebastian Raschka, Leslie A. Kuhn, Anne M. Scott, and Weiming Li (2018) *Computational Drug Discovery and Design: Automated Inference of Chemical Group Discriminants of Biological Activity from Virtual Screening Data*. Springer. ISBN: 978-1-4939-7755-0.
- Mike Driscoll (ed.) (2018). *Python Interviews: Discussions with prolific programmers*. Packt Publishing Ltd. ISBN: 978-1-7883-9908-1.



## Selected Oral Presentations and Talks

---

- *Designing Generative Adversarial Networks for Privacy-enhanced Face Recognition*. 14th Int. Conf. On Human-System Interaction (IEEE, IES). Virtual. July 8th, 2021.
- *An Introduction to Generative Adversarial Networks in Computer Vision*. *International Summer School on Deep Learning* (Virtual). Gdansk University of Technology, Poland. July 6th, 2021.
- *Modern Machine Learning: An Introduction to the Latest Techniques*. Invited talk at the Chan Zuckerberg Initiative's Computational Seed Network conference.. April 6th, 2021.
- *Machine Learning in Python: Recent Trends, Technologies, and Challenges*. ODSC East. March 31st, 2021.
- *UW-Madison Statistics Teaching Expo 2020 – Panelist*
- *Deep Learning with Python in 2020: Recent Technological Advances and Research Trends (invited; canceled due to COVID-19)*. July 2020. International Summer School on Deep Learning. Gdansk University of Technology, Poland.
- *Deep Generative Models for Image Generation (invited; canceled due to COVID-19)*. July 2020. International Summer School on Deep Learning. Gdansk University of Technology, Poland
- *Machine Learning with Python in 2020: Our Technology and Challenges*. Port Harcourt School of AI Nigeria (<https://phcschoolofai.org>). April 2020. Port Harcourt, Nigeria. Remote.
- *Machine Learning with Python in 2020: Our Technology and Challenges*. Open Data Science Conference (ODSC East 2020). April 2020. Boston.
- *Machine Learning and Deep Learning: Technologies, Challenges, and Hot Research Areas*. March 18. 3M. Minneapolis (Remote)
- *Predicting and Hiding Personal Information from Face Images using Deep Learning (invited)*. October 2019. ML4MI Initiative Seminar Series, University of Wisconsin Madison.
- *Predicting soft-biometric attributes from face images with deep learning and PyTorch (invited)*. October 2019. UW-Madison Data Science Club.
- *Machine-Learning & AI-based Approaches for GPCR Bioactive Ligand Discovery (invited)*. September 2019. Cambridge Healthtech Institute's 14th Annual GPCR-Based Drug Discovery – Discovery on Target Conference
- *Predicting and Hiding Personal Information from Face Images (invited)*. August 2019. Big Data Madison
- *Convolutional Neural Networks for Predicting and Hiding Personal Traits from Face Images (invited)*. July 2019. International Summer School on Deep Learning. Gdansk University of Technology, Poland

- *Training Neural Networks for Ordinal Regression (invited)*. July 2019. International Summer School on Deep Learning. Gdansk University of Technology, Poland
- *Predicting and Hiding Personal Traits from Face Images using Convolutional Neural Networks (invited)*. Apr 2019. Biology of Brain and Behavior (BBB) Seminar Series. Dept. of Psychology at UW-Madison
- *Discovering patterns of molecular recognition: Protein-ligand interfaces are polarized and favor hydrogen bond donors on the protein side (invited)*. Mar 2019. CIBM series at UW-Madison
- *Imparting privacy to face images: designing semi-adversarial neural networks for multi-objective function optimization (invited)*. Apr 2018. Applied Machine Learning Conference 2018. Charlottesville, VA.
- *Multi-objective optimization with semi-adversarial neural networks (invited)*. University of Wisconsin-Madison Statistics Seminar Series. Feb 2018. Madison, WI.
- *Machine Learning with Python (invited)*. MSU Data Science. Feb 2018. Michigan State University, East Lansing, MI.
- *Building hypothesis-driven virtual screening pipelines for millions of molecules*. ODSC West. Nov 2017. San Francisco, CA.
- *Introduction to Deep Learning with TensorFlow (invited)*. PyData Ann Arbor. Aug 2017. Ann Arbor, MI.
- *Efficient large-scale virtual screening with molecular data frames*. SciPy 2017 Lightning Talk. Jul 2017. Austin, TX.
- *Binding pose prediction via rigidity analysis and large-scale pheromone antagonist discovery for invasive species control (invited)*. MSU Computational and Systems Biology Club. Mar 2017. East Lansing, MI.
- *Machine Learning and Performance Evaluation – Overcoming the Selection Bias (invited)*. DataPhilly. Nov 2016. Philadelphia, PA.
- *Detecting the Native Ligand Orientation by Interfacial Rigidity (invited)*. BMB Departmental Retreat. Oct 2016. Okemos, MI.
- *Getting Started with Data Science (invited)*. MSU DataScience. Sep 2016. East Lansing, MI.
- *Learning Scikit-Learn (invited)*. PyData 2016. Aug 2016. Chicago, IL.
- *Machine Learning with Scikit-Learn*. SciPy 2016. Jul 2016. Austin, TX.
- *A Novel Approach to Protein-Ligand Binding Mode Prediction by Rigidity Analysis Using Graph Theory (invited)*. Biomolecular Science Gateway. Feb 2016. East Lansing, MI.
- *Practical Data Science. An Introduction to Supervised Machine Learning and Pattern Classification (invited)*. MSU NextGen Bioinformatics. Feb 2015. East Lansing, MI.

## Poster Presentations

---

- Benjamin Kaufman and Sebastian Raschka. *A Study of Graph Neural Networks for Computing Strain Energy of Molecular Conformer Ensembles* CIBM and BDS Annual Retreat 2019, University of Wisconsin-Madison
- V. Mirjalili, S. Raschka, A. Ross. *Semi-Adversarial Neural Networks: A New Paradigm for Imparting Privacy to Face Images*. Engineering Research Symposium, Michigan State University, March 2018 (Best Poster Award)
- V. Mirjalili, S. Raschka, A. Namboodiri, A. Ross. *Semi Adversarial Networks: Retaining Biometric Utility of Face Images while Perturbing Gender*. 11th IAPR International Conference on Biometrics (ICB 2018), Gold Coast, Queensland, Australia, February 20-23, 2018
- S. Raschka, N. Liu, S. Gunturu, A.M. Scott, M. Huertas, W. Li, and L.A. Kuhn. *Screenlamp: A Software Framework for Hypothesis-Driven Ligand Discovery Based on Virtual Screening and Machine Learning*. Great Lakes Bioinformatics and Canadian Computational Biology Conference 2016, University of Toronto, May 16-19, 2016
- A. Scott, M. Huertas, S. Raschka, N. Liu, L.A. Kuhn, and W. Li. *Regulation of Pheromone Induced Responses with Behavioral Antagonists for Invasive Sea Lamprey (*Petromyzon marinus* L.) Control*. Behaviour 2015 conference, Cairns, Australia, August 8, 2015
- S. Raschka, N. Liu, S. Gunturu, A.M. Scott, M. Huertas, W. Li, and L.A. Kuhn. *A Hypothesis-Driven Virtual Screening Methodology for Structure-Based Inhibitor Discovery*. Quantitative Biology Symposium, Michigan State University, East Lansing, August 18, 2014
- A. Scott, M. Huertas, N. Liu, S. Gunturu, S. Raschka, L. A. Kuhn, and W. Li. *Identification of pheromone receptor antagonists for invasive sea lamprey control*. International Congress on the Biology of Fish, Heriot-Watt University, UK, 3-7 August, 2014

## Professional Service and Journals

---

- Moderator for the machine learning (CS.LG) part of CoRR in arXiv (2019-present, URL: <https://arxiv.org/corr/subjectclasses>)
- "Machine Learning with Python," a special issue of Information (ISSN 2078-2489). This special issue belongs to the section "Artificial Intelligence" (Guest Editor, URL: [https://www.mdpi.com/journal/information/special\\_issues/ML\\_Python](https://www.mdpi.com/journal/information/special_issues/ML_Python))

## Peer Reviewing

---

- 2021:
  - International Conference on Machine Learning (ICML)
  - Nature Machine Intelligence
  - PeerJ Computer Science,

- PLOS ONE
  - Nature Scientific Reports
  - Conference on Computer Vision and Pattern Recognition (CVPR)
  - Journal of Chemical Information and Modeling (JCIM)
  - International Conference on Computer Vision (ICCV)
  - The Journal of Open Source Software (JOSS)
- **2020:**
- Winter Conference on Applications of Computer Vision (WACV '21)
  - PLOS Computational Biology
  - Nature Scientific Reports
  - BioScience (Oxford Academic)
  - Journal of Machine Learning Research (JMLR)
  - Wiley ChemistrySelect
  - UW Institute for Clinical and Translational Research (UW ICTR) 2020 pilot grant proposals
  - Advances and Applications in Mathematical Sciences, 2020
  - "Research Funding from American Family Insurance" competition 2020 grant proposals
  - MDPI Mathematics
  - International Conference on Machine Learning (ICML)
- **2019:**
- IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS 2019)
  - Journal of Machine Learning Research (JMLR)
  - Notices of the American Mathematics Society (AMS)
  - Journal of Applied Statistics (CJAS)
  - Journal of Chemical Information and Modeling (JCIM)
- **2018:**
- Scientific Reports, Nature Publishing Group
  - Bioinformatics, Oxford University Press
  - The Journal of Open Source Software (JOSS)
- **2017:**
- Biological Invasions (BINV), Springer

## Open Source Projects and Contributions

---

- Creator and maintainer of the Python data science library mlxtend (<https://github.com/rasbt/mlxtend>)
- Creator and maintainer of Screenlamp, A Python toolkit for facilitating hypothesis-driven, ligand-based screening of large molecule libraries (<https://github.com/psa-lab/screenlamp>)
- Creator and maintainer of molecular data frame library BioPandas (<http://rasbt.github.io/biopandas>)
- Creator and maintainer of SiteInterlock, a novel approach to pose selection in protein-ligand docking based on graph theory (<https://github.com/psa-lab/siteinterlock>)
- Creator and maintainer of a collection tutorials and examples for solving and understanding machine learning and pattern classification tasks ([https://github.com/rasbt/pattern\\_classification](https://github.com/rasbt/pattern_classification))
- Author of Python Machine Learning with code material under a open source license on GitHub (1st Ed.: <https://github.com/rasbt/python-machine-learning-book>, 2nd Ed.: <https://github.com/rasbt/python-machine-learning-book-2nd-edition>)
- Author of "Introduction to Artificial Neural Networks and Deep Learning: A Practical Guide with Applications in Python" with code material under a open source license on GitHub (in progress, <https://github.com/rasbt/deep-learning-book>)
- Creator and maintainer of the "Automated Inference of Chemical Discriminants of Biological Activity" protocols (<https://github.com/psa-lab/predicting-activity-by-machine-learning>)
- Creator and maintainer of "Algorithms in IPython Notebooks," A repository with IPython notebooks of algorithms implemented in Python ([https://github.com/rasbt/algorithms\\_in\\_ipython\\_notebooks](https://github.com/rasbt/algorithms_in_ipython_notebooks))
- Creator and maintainer of the Python Progress Indicator utility package, which is being shipped with Debian, a Unix-like computer operating system (<https://github.com/rasbt/pyprind>)
- Journal of Open Source Software reviewer
- Contributor to the Python machine learning library Scikit-learn scikit-learn
- Contributor to the deep learning library TensorFlow
- Member of the conda-forge organization, which provides recipes, build infrastructure, and distributions for the conda package manager
- Member of the Python Software Foundation
- An extended open source portfolio is available on GitHub: <https://github.com/rasbt>

## Honors, Awards, and Research Support

---

- NVIDIA Travel Award (2020)
- Google Cloud Education Grant (2019)
- NVIDIA GPU Grant (2018)
- Best Poster Award, Michigan State University, Engineering Research Symposium 2018
- Best Paper Award, International Conference on Biometrics (ICB 2018)
- MSU Departmental Outstanding Graduate Student Award (2016-2017)
- ACM Computing Reviews' Best of 2016 (Python Machine Learning book)
- Quora Top Writer 2016 (Machine Learning)
- OpenIntro Statistics outstanding student paper (2011)

## Teaching Experience

---

- Spring 2021** STAT453 SS 2021: Introduction to Deep Learning and Generative Models  
*Principal Instructor*  
4-month course. Lecture design, teaching in-class sessions. Homework, exam, and course project design. 72 students.  
Instructor evaluation score: 4.76/5.00.  
Course evaluation score: 4.79/5.00.
- Spring 2021** MinneMUDAC  
*Advisor*  
Advised 1 team who participated in the MinneMUDAC Student Data Science Challenge, <https://minneanalytics.org/minnemudac/>.
- Fall 2020** STAT451 FS 2020: Introduction to Machine Learning and Statistical Pattern Classification  
*Principal Instructor*  
4-month course. Lecture design, teaching in-class sessions. Homework, exam, and course project design. 65 students.  
Instructor evaluation score: 4.82/5.00.  
Course evaluation score: 4.61/5.00.

- Fall 2020** STAT571 FS 2020: Statistical Methods for Bioscience I  
*Principal Instructor*  
4-month course. Teaching in-class sessions. Homework and exam. 140 students.  
Instructor evaluation score: 4.24/5.00.  
Course evaluation score: 4.10/5.00.
- Spring 2020** MUDAC  
*Advisor*  
Advised 1 team who participated in the Midwest Data Analytics Competition (MUDAC), <http://www.mudac.org>.
- Spring 2020** STAT453 SS 2020: Introduction to Deep Learning and Generative Models  
*Principal Instructor*  
4-month course. Lecture design, teaching in-class sessions. Homework, exam, and course project design. 62 students.  
Course evaluation score: 4.75/5.00.
- Fall 2019** STAT479 FS 2019: Machine Learning  
*Principal Instructor*  
4-month course. Lecture design, teaching in-class sessions. Homework, exam, and course project design. 72 students.  
Course evaluation score: 4.57/5.00.
- July 2019** International Summer School on Deep Learning  
*Keynote Speaker & Instructor*  
Keynote talk and tutorial session on ordinal regression with deep learning. More than 100 attendees from more than 25 countries. Summer school website: <http://2019.dl-lab.eu>
- Spring 2019** MUDAC  
*Advisor*  
Advised three teams who participated in the Midwest Data Analytics Competition (MUDAC), <http://www.mudac.org>.
- Spring 2019** STAT479 SS 2020: Machine Learning  
*Principal Instructor*  
4-month course. Lecture design, teaching in-class sessions. Homework, exam, and course project design. 72 students.

Course evaluation score: 4.32/5.00.

- Fall 2018** STAT479 FS 2018: Machine Learning  
*Principal Instructor*  
4-month course. Lecture design, teaching in-class sessions. Homework, exam, and course project design. 72 students.  
Course evaluation score: 4.76/5.00.
- Feb 2018** Machine Learning Workshop  
*Principal Instructor*  
Machine learning workshop at Michigan State University, MSU Data Science Club.  
Approx. 100 attendees.
- Aug 2012-2017** Project Design  
*Co-Supervisor*  
Composed projects for undergraduate and graduate students in the Protein Structure and Analysis Laboratory and trained students in computational and statistical subjects.
- Aug 2016** Machine learning with scikit-learn at SciPy 2016, Austin, Texas  
*Instructor*  
Taught an 8-hour workshop on machine learning and scikit-learn (approx. 100 participants).
- May 2014-  
Jul 2014** BMB 401: Comprehensive Biochemistry, Michigan State University  
*Graduate Teaching Assistant*  
Co-supervised the distance learning based section of the course that required the extensive use of computational teaching technology to interact with the students and provide course materials (approx. 300 students).
- May 2013** Software Carpentry Instructor Training, Toronto, Canada  
Attended the one-week training for Software Carpentry instructors.
- Jun 2010-  
Aug 2010** Bio 250: Genetics, Heinrich-Heine Universitaet Duesseldorf  
*Student Teaching Assistant*



Co-supervised laboratory sessions and designed evaluated weekly written tests to help students with the course material and preparation for the exam (approx. 200 students).

## Clubs/Extracurricular Activities

---

**Aug 2016-  
Jan 2018** MSU Data Science Club  
*Member and speaker*

Advised MSU students in applying data science techniques to research problems.

**Oct 2013-  
Apr 2014** Council of Graduate Students, Michigan State University  
*Departmental representative*

Participated in council meetings and discussions, representing the Department of Biochemistry and Molecular Biology.

## Recent and Selected Community Outreach

---

- Co-Sponsor of the Women in Coding and Data Science (WiCDS) Data Science and Machine Learning Blogathon, 2020.
- Guest lecture on *Machine Learning in Python at the Port Harcourt School of AI* in Nigeria, 2020

## Students

---

Current Graduate Students

- Jitian Zhao, Ph.D. student
- Ben Kaufman, Ph.D. student
- Xintong Shi, M.Sc. student

Former Graduate Students

- Dr. Zhongjie Yu, Thesis: "Few-Shot Learning: Contributions to Deep Learning With Limited Data". After graduation: Research Scientist at Wyze Labs.

## Committees Served

---

- Dep. of Statistics IT Position Search Committee, 2020-2021
- VISP & M.Sc. of Data Science Committee (Dep. of Statistics), 2020-21
- Computing & Web Committee (Dep. of Statistics), 2020-21
- VISP & M.Sc. of Data Science Committee (Dep. of Statistics), 2019-20
- Undergrad Studies Committee (Dep. of Statistics), 2019-20
- Computing & Web Committee (Dep. of Statistics), 2018-19
- Undergrad Studies Committee (Dep. of Statistics), 2018-19
- Ph.D. committees: Bowen Hu (2020-present); Liam Johnston (2020-present); Moayad Alnammi (2019-2021; successfully defended); Tony Meger (2018-present); Jiacheng Cheng (2018-2020, successfully defended)

## Selected Journalism Contributions and Interviews

---

- The African Data Scientist Podcast, 2020 (by Port Harcourt School of AI, Nigeria): *#6: On how Africa can thrive in the age of AI and Data Science, and AI and Data Science in the African educational ecosystem*
- Deep Tech Musings podcast, 2020. Drug Discovery with ML/AI
- Podcast `__Init__` (Pythonpodcast.com), 2020: *Teaching Python Machine Learning - Episode 260 – Podcast Interview*
- Chai Time Data Science, 2020: *Statistics, Open Source & ML Research – Podcast Interview*
- Wisconsin State Journal, 2019: *Advancing machine learning while protecting privacy*
- DataCamp, 2018: *Biology and Deep Learning – Podcast Interview*
- Partially Derivative Podcast Interview, 2017: *Model Evaluation with Sebastian Raschka*
- Computer Business Reviews, 2017: *The Future of Machine Learning and Data Science*
- KDnuggets, 2016: *Top 2016 KDnuggets Stories*
- KDnuggets, 2016: *Data Preparation Tips, Tricks, and Tools: An Interview with the Insiders*
- Huffington Post, 2016: *The Impact of Machine Learning on Healthcare*
- Forbes, 2016: *What Do Computational Biologists Do?*

- Hacker Bits print magazine, 2016: *Everyone is talking about deep learning these days, but what the heck is it?*
- Open Data Science, 2016: *Trifecta: Python, Machine Learning, and Dueling Languages A.I.*
- Open Data Science, 2016: *Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning*
- Data Science at Home Podcast Interview, 2016: *30 min with data scientist Sebastian Raschka*
- Becoming a Data Scientist Podcast, 2016: *Interview with Sebastian Raschka about how he became a data scientist, his current research, and about his book*

## Selected Press Coverage

---

- AI Newsletter, 2021: *The 100 Most Influential People In AI – 2021 Edition*
- Pythonbooks.org, 2020: *Best Python Book of The Month*
- Big Think, 2018: *The 10 best books about A.I.*
- Onanalytica, 2018: *Top Influential Developers in AI*
- Book Scrolling, 2018: *The Best Data Science Books of All-time*
- A.I. & Optimization, 2016: *Top 22 Python Programming Books*
- KDnuggets, 2016: *Top 10 Essential Books for the Data Enthusiast*
- KDnuggets, 2016: *5 Machine Learning Projects You Can No Longer Overlook*
- Techopedia, 2016: *Artificial Intelligence: Experts To Follow on Twitter*
- InformationWeek, 2016: *10 Big Data Books To Boost Your Career*
- OpenSource.org, 2015: *Publisher's picks: 29 open source books for 2015*
- Datawerq, 2015: *50 Data leaders to follow on Twitter*
- Onanalytica, 2015: *Artificial Intelligence & Machine Learning: Top 100 Influencers and Brands*
- Analytics Vidhya, 2015: *Top Data Scientists to follow on GitHub*