

Scott Lundberg

explainable artificial intelligence for science and medicine

Contact

scott@scottlundberg.com
http://scottlundberg.com

Languages

Python/C/C++/Julia
Javascript
Swift/Java/Perl/R
Bash/PHP/SQL/Matlab

Software

Numpy/Jupyter/Pandas
Scikit-learn/Keras
XGBoost/LightGBM
TensorFlow/PyTorch
Angular/React
ES/Solr/Mongo/MySQL
AWS/Google Cloud

My research focuses on explainable artificial intelligence and its application to problems in medicine and healthcare. This has led to the development of broadly applicable methods and tools for interpreting complex machine learning models.

Education

University of Washington, Seattle, WA
Doctor of Philosophy, Computer Science & Engineering; Spring 2019
Advisor: Prof. Su-In Lee; Emphasis: Interpretable machine learning in medicine

Colorado State University, Fort Collins, CO
Master of Science, Computer Science; May 2008
Advisor: Prof. Ross McConnell; Emphasis: Graph theory

Bachelor of Science, Computer Science; December 2005; *summa cum laude*

Open Source Software: SHAP

SHAP is an implementation of the explainable AI methods we have developed. It is used by medical providers, logistics companies, banks, insurance companies, NBA teams, university researchers, cloud providers, startups, and major tech companies. It includes custom algorithms to interpret models from XGBoost, LightGBM, CatBoost, scikit-learn, TensorFlow, and PyTorch. It has 8,000+ GitHub stars (since Jan 2018); and 1,200+ forks: github.com/slundberg/shap

Research/Work Experience

- | | |
|--------------|---|
| 2019-present | Senior Researcher at Microsoft Research. |
| 2013-2019 | Ph.D. Graduate Student at the University of Washington. Led several research efforts, mentored younger graduate students, served as a teaching assistant for three courses, taught lab recitations, and was a guest lecturer in classes. |
| 2018-2019 | Interpretable Machine Learning Consulting. Giving workshops, along with implementation and guidance on the use of explainable AI techniques. |
| 2008-2013 | Research Scientist at Numerica Corporation. Helped design, and launch the Lumen data analysis product that now serves police departments in many states. Served as the principal investigator for our Chemical/Biological detection program (with a clearance). Helped develop patented GIS data compression. Worked on viral protein identification, radar tracking, novelty detection, and sensor data fusion. |
| 2006-2008 | Research/Teaching Assistant at Colorado State University. Worked on improving the complexity of various graph theory algorithms. My teaching assistant role covered classes in algorithms, databases, and computer networks. |
| 2006-2007 | Internship at Numerica Corporation. Designed and implemented improvements to state-of-the-art assignment solver algorithms. |
| 2003-2008 | Web Hosting and Automated Design Business. Sole proprietorship based largely around what was then the beginnings of modern single-page web applications. |
| 2001-2003 | Internship at StarFire Enterprises. Worked as a web developer and system administrator for UNIX based systems in high school. |

Honors/Awards

- 2019 ODSC Open Data Science Award
- 2014-2019 National Science Foundation Graduate Research Fellowship
- 2018 Paper selected as the cover article for Nature Biomedical Engineering
- 2018 Best of Clinical Research Abstract (UW's annual anesthesia research day)
- 2017 Paper selected for an oral presentation at NeurIPS (top 1% of submissions)
- 2017 Madrona Prize (based on the commercial potential of our research)
- 2017 AARP Foundation Health Innovation Prize (based on our potential health startup)
- 2016 Best paper at the NeurIPS workshop on interpretable machine learning
- 2009 Outstanding Platform Presentation at the Chem/Bio Defense S&T Conference
- 2005 Patricia R. Mohilner Memorial Scholarship

Selected Presentations

- 2019 Invited talk at the U.S. Consumer Financial Protection Bureau
- 2018 Invited keynote talk at the Interpretability of Machine Intelligence in Medical Image Computing workshop at MICCAI (the International Conference On Medical Image Computing & Computer Assisted Intervention)
- 2017 Oral Presentation at NeurIPS (two track conference with 8k attendees)
- 2017 Open Data Science Conference Talk on Interpretable Machine Learning
- 2017-present Various invited talks at companies and organizations such as Tableau, Uber, Two Sigma, Harborview Injury Prevention & Research Center, and others.

Publications/Patents

- 2020 Explaining Quantitative Measures of Fairness
S. Lundberg; Fair & Responsible AI Workshop @ CHI2020.
- 2019 Explainable AI for Trees: From Local Explanations to Global Understanding (*cover article*)
S. Lundberg, G. Erion, H. Chen, A. DeGrave, J. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, S. Lee; Nature Machine Intelligence.
- 2019 AIControl: replacing matched control experiments with machine learning improves ChIP-seq peak identification
N. Hiranuma, S. Lundberg, S. Lee; Nucleic Acids Research.
- 2018 Explainable machine-learning predictions for the prevention of hypoxaemia during surgery (*cover article*)
S. Lundberg, B. Nair, M. Vavilala, M. Horibe, M. Eisses, T. Adams, D. Liston, D. Low, S. Newman, J. Kim, and S. Lee; Nature Biomedical Engineering.
- 2018 Consistent Individualized Feature Attribution for Tree Ensembles
S. Lundberg, G. Erion, S. Lee; arXiv preprint.
- 2018 A machine learning approach to integrate big data for precision medicine in acute myeloid leukemia
S. Lee, S. Celik, B. Logsdon, S. Lundberg, T. Martins, V. Oehler, E. Estey, C. Miller, S. Chien, J. Dai, and A. Saxena; Nature communications.

- 2017 A unified approach to interpreting model predictions (*100+ citations in 2018*)
selected for oral presentation (top 1% of all submissions)
S. Lundberg, S. Lee; NeurIPS.
- 2017 Anesthesiologist-level forecasting of hypoxemia with only SpO2 data using deep learning
G. Erion, H. Chen, S. Lundberg, S. Lee; NeurIPS Workshop ML4H: ML for Health.
- 2017 Hybrid Gradient Boosting Trees and Neural Networks for Forecasting Operating Room Data
H. Chen, S. Lundberg, and S. Lee; NeurIPS Workshop: ML for Health.
- 2016 An unexpected unity among methods for interpretable model predictions (*selected for best paper award*)
S. Lundberg and S. Lee; NeurIPS Workshop: Interpretable Machine Learning for Complex Systems.
- 2016 Cloud Control: Leveraging many public ChIP-seq control experiments to better remove background noise
N. Hiranuma, S. Lundberg, and S. Lee; ACM-BCB.
- 2016 Learning the human chromatin network using all ENCODE ChIP-seq datasets (F1000Prime recommended)
S. Lundberg, W. Tu, B. Raught, L. Penn, M. Hoffman, S. Lee; Genome Biology.
- 2014 Method for Lossy Compression of Point Clouds with Pointwise Error Constraints
R. Nong, R. Paffenroth, S. Lundberg, and W. Leed; U.S. Patent.
- 2012 Detecting Clustered Chem/Bio Signals in Noisy Sensor Feeds Using Adaptive Fusion
S. Lundberg, C. Calderon, and R. Paffenroth; SPIE Conference Series.
- 2011 A data-driven approach for processing heterogeneous categorical sensor signals
C. Calderon, A. Jones, S. Lundberg, and R. Paffenroth; SPIE Conference Series.
- 2010 $O(m \log n)$ split decomposition of strongly-connected graphs
B. Joeris, S. Lundberg, and R. McConnell; Discrete Applied Mathematics.
(alphabetical author order)
- 2010 Analysis of CBRN sensor fusion methods
S. Lundberg, R. Paffenroth, and J. Yosinski; Information Fusion (FUSION) IEEE.
- 2010 An implicit representation of chordal comparability graphs in linear time
A. Curtis, C. Izurieta, B. Joeris, S. Lundberg, and R. McConnell; Discrete Applied Mathematics. (*alphabetical author order*)
- 2010 Algorithms for Distributed Chemical Sensor Fusion
S. Lundberg, R. Paffenroth, and J. Yosinski; Proceedings of SPIE.
- 2008 $O(m \log n)$ split decomposition of directed graphs
S. Lundberg; Master's thesis, Colorado State University.
- 2008 Top down image segmentation using congealing and graph-cut
D. Moore, J. Stevens, S. Lundberg, and B. Draper; ICPR, IEEE.