Gerrit J.J. van den Burg, PhD

London, UK | Email: gertjanvandenburg@gmail.com | Web: gertjanvandenburg.com

Summary

I am a research scientist with 8+ years of experience in academic and industry machine learning research. I enjoy working on complex, large-scale problems that can have a positive real-world impact. I have extensive experience with machine learning modeling, algorithm design, and software engineering. My goal is to use my skills and expertise to address ongoing challenges in machine learning and AI.

Research Experience

Applied Scientist II — Amazon Alexa, UK 2021–Present

- Developing methods to improve the accuracy of speech recognition systems

Postdoctoral Researcher — The Alan Turing Institute, UK 2018–2021

- Introduced a memorization score for probabilistic deep generative models and showed that neural networks can remember part of their input data, which has important implications for data privacy
- Created a method for structure detection in textual data files that improved on the Python builtin method by 21%. Developed this into a Python package that has received over 1M downloads
- Developed a robust Bayesian matrix factorization algorithm for time series modeling and forecasting that improved imputation error up to 60% while maintaining competitive runtime
- Established the first benchmark dataset for change point detection on general real-world time series and determined the best performing methods, with consequences for research and practice

Doctoral Researcher — Erasmus University Rotterdam, NL 2012–2018

- Formulated a generalized multiclass SVM classifier that significantly improved on 4 existing alternatives in terms of accuracy and runtime on several datasets
- Designed an optimization algorithm for non-convex sparse regularization in regression problems that outperformed existing methods on parameter estimation and prediction accuracy
- Improved an estimator for measuring the difficulty of classification problems by more than 30% and leveraged this into a novel hierarchical classification method using SVMs

Undergraduate research — Delft University of Technology, NL 2011–2012

- Studied physics models for 1D contact processes using Monte Carlo simulations on the GPU
- Implemented imaging algorithms for protein measurements with applications to dementia research

Education

- Ph.D. in Machine Learning, Erasmus University Rotterdam, NL 2018
- Visiting researcher with prof. Alfred Hero (4 mo.), University of Michigan, USA 2016
- Visiting student with prof. Patrick Groenen (1 mo.), Stanford University, USA 2014
- M.Sc. Econometrics, Erasmus University Rotterdam, NL 2012
- M.Sc. Applied Physics, Delft University of Technology, NL 2012
- B.Sc. Applied Physics, Delft University of Technology, NL 2009

Awards & Grants

- **Best Reviewer Award**, Neural Information Processing Systems Conference 2020
- **Top 33% Reviewer**, International Conference on Machine Learning 2020
- **Best Reviewer Award**, Neural Information Processing Systems Conference 2019
- **Top Educator Award**, Erasmus School of Economics 2016
- **Research Grant**, Erasmus Research Institute of Management 2016

Skills

- Author of twelve Python packages and three R packages (1M+ downloads combined)
- Experienced in modern machine learning paradigms (deep learning, variational autoencoders, Gaussian processes, convolutional neural networks, etc.)
- Proficient in software engineering tools and practices (Git, Scrum/Agile, continuous integration)
- Programming languages: Python, C, R, MATLAB, Javascript
- Software: PyTorch, TensorFlow, Scikit-learn, NumPy, Pandas, Cython, Docker, Linux, Make, Bash
- Languages (fluent): English, Dutch

Publications

International Conferences

- **G.J.J. van den Burg**, C.K.I. Williams. On Memorization in Probabilistic Deep Generative Models. Advances in Neural Information Processing Systems (NeurIPS), 2021.
- Ö.D. Akyildiz*, **G.J.J. van den Burg***, T. Damoulas, M.J.F. Steel. Probabilistic Sequential Matrix Factorization. International Conference on Artificial Intelligence and Statistics (AISTATS), 2021. (* = joint first author)

Journals

- **G.J.J. van den Burg**, A. Nazábal, C. Sutton. Wrangling Messy CSV Files by Detecting Row and Type Patterns. *Data Mining and Knowledge Discovery*, 33(6):1799–1820, 2019.
- **G.J.J. van den Burg** and P.J.F. Groenen. GenSVM: A Generalized Multiclass Support Vector Machine. *Journal of Machine Learning Research*, 17(225):1–42, 2016.

Book chapters

- **G.J.J. van den Burg**, Reproducible Research with Make. In "The Turing Way – A Handbook for Reproducible Data Science", 2019.

Preprints

- **G.J.J. van den Burg** and C.K.I. Williams. An Evaluation of Change Point Detection Algorithms. *arXiv:2003.06222*, 2020.
- **G.J.J. van den Burg** and A.O. Hero. Fast Meta-Learning for Adaptive Hierarchical Classifier Design. *arXiv:1711.03512*, 2017.
- **G.J.J. van den Burg**, P.J.F. Groenen, A. Alfons. SparseStep: Approximating the Counting Norm for Sparse Regularization. *arXiv:1701.06967*, 2017.