David Bosch, PhD

Gothenburg, Sweden | davbosch01@gmail.com | +46 76 626 04 41 | Google Scholar | GitHub

Professional Summary

Machine Learning Researcher and Data Scientist with a PhD in Computer Science and Engineering (Chalmers University of Technology). Experienced in large-scale model analysis, statistical learning theory, and NLP systems.

Education

Chalmers University of Technology, Gothenburg, Sweden *Jun 2020 – Oct 2025*
PhD in Computer Science and Engineering

- Research focused on the asymptotic analysis of machine learning models; the analysis of the statistical properties of ML models, in the regime where the number of data points and model parameters grow asymptotically large.
- Extended the literature on comparison theorems and universality theorems applied to ML models, by developing novel theorems to analyze a larger class of models, with a focus on using statistics and statistical physics frameworks as tools of analysis.
- Numerically validated theoretical results by means of simulations in Python, using tools such as NumPy, SciPy, and PyTorch.
- Published in top conferences such as AISTATS and COLT, as well as journals such as IEEE transactions on information theory.

University of Groningen, Netherlands *Sep 2018 – Jun 2020*
MSc in Theoretical Physics

- Wrote high performance Lattice Quantum Chromodynamics simulations in C to analyze the statistical properties of gluons, a theoretical particle predicted by the standard model of particle physics.

University of Groningen, Netherlands *Sep 2015 – Jun 2018*
BSc in Physics

Experience

Data Scientist *Kwantum Analytics (Remote) — Sep 2021 – Oct 2024*

- Designed and deployed end-to-end Natural Language Processing (NLP) pipelines to analyze open-ended responses with more than 10,000 responses, improving insight extraction from unstructured text data, and reducing analysis time by more than 100%.
- Implemented embedding-based analysis using BERT and Llama 3 models for semantic clustering, topic discovery, and downstream analytics.
- Built Bayesian network models for causal key-driver analysis of marketing KPIs, enabling data-driven strategic decisions.
- Developed a prototype educational chatbot using Retrieval-Augmented Generation (RAG) over curriculum documents using Microsoft Azure, Python, and React.

Lecturer and Teaching Assistant *Chalmers University of Technology — Sep 2020 – Mar 2025*

- Gave lectures and led tutorials on Python, Machine Learning, and Linear Programming courses with more than 80+ master students.
- Supported course design, assignment development, and direct mentoring of student projects.

Selected Publications

- **Precise Asymptotic Analysis of Deep Random Feature Models**, *COLT 2023*.
- **Random Features Model with General Convex Regularization**, *AISTATS 2023*.
- **A Novel Gaussian Min-Max Theorem and its Applications**, *IEEE Trans. Information Theory, 2025*.
- Full list: Google Scholar.

Technical Skills

Programming: Python, Rust, C
ML Frameworks: PyTorch, TensorFlow, Scikit-learn, Hugging Face, NumPy, SciPy
Tools: Git, SQL, Microsoft Azure
Core Strengths: Asymptotic ML analysis, Statistical Learning Theory, Statistical physics for ML
Languages: English (fluent), Dutch (fluent)