Soufiane Hayou

Simons Institute for the Theory of Computing, UC Berkeley
✉ hayou@berkeley.edu
🖥 https://www.soufianehayou.com/

Last updated: 6/8/2024

Academic Appointments

Simons Institute for the Theory of Computing, UC Berkeley **Berkeley, United States**
Researcher *2023 – Current*
- Member of the Collaboration on the Theoretical Foundations of Deep Learning. Hosted by Bin Yu and Peter Bartlett.
- Investigating statistical properties of large scale neural networks (large width, large depth, etc).
- Developing scalable efficient learning methods for the pretraining and finetuning of large models.

National University of Singapore **Singapore**
Peng Tsu Ann Assistant Professor of Mathematics (On Leave) *2021 – Current*
- Led multiple research projects on neural network scaling.
- Taught core ML courses at the graduate level: Principles of Machine Learning (DSA5105, 100+ students), Advanced Topics in Machine Learning (DSA5202).

University of Oxford **Oxford, United Kingdom**
Class Tutor and Teaching Assistant *2018 – 2020*
- Advanced simulation methods (Undergraduate level).
- MCMC and Diffusions methods (Graduate level).

Education

University of Oxford **Oxford, United Kingdom**
PhD in Machine Learning & Statistics *2017 – 2021*
Advised by Arnaud Doucet and Judith Rousseau.
- Thesis: Wide Deep Neural Networks.

Pierre et Marie Curie University (Paris VI) **Paris, France**
MSc. in Probability and Financial Mathematics with Distinction. *2016 – 2017*

Ecole Polytechnique **Paris, France**
BSc and MSc in Applied Mathematics and Engineering Diploma with Distinction. *2013 – 2017*

Research Interests

Theory and Methods:
- Large scale neural networks: stability, learning dynamics, etc.
- Efficient Learning at Scale: efficient algorithms for the pretraining and finetuning of large models.
- Algorithms for generating and filtering synthetic data.
- Model and data compression (pruning).
- Different aspects of Large Language Models (LLMs), particularly Hallucination and Privacy.
- Retrieval-Augmented Generation (RAG) (see EduBoost project below).

Applications:
- AI applications in Life Sciences, Decision Science, and Education (see EduBoost project below).
- Trustworthy AI Agents.

Projects' highlights:
- Inventor of Stable ResNet: A modified architecture enhancing stability in very deep neural networks, contributing to improved training of such models.

- Co-Inventor of Depth-μP: An innovative technique for scaling the number of layers in neural networks (depth version of μP).
- Inventor of LoRA+: An advanced method for efficient low-rank adaptation of large models, currently integrated into HuggingFace's PEFT package and and LLamaFactory.
- Creator of "EduBoost": An AI teaching assistant designed to support course instructors. It digests course content and provides students with answers to their questions.

Publications and Preprints

✧ Yoonsoo Nam, Chris Mingard, Seok Hyeong Lee, **Soufiane Hayou**, Ard Louis. "Visualising Feature Learning in Deep Neural Networks by Diagonalizing the Forward Feature Map". (Submitted, 2024).

✧ **Soufiane Hayou**, Nikhil Gosh, Bin Yu. "The Impact of Initialization on LoRA Finetuning Dynamics". In: *Neural Information Processing Systems* (NeurIPS 2024).

✧ **Soufiane Hayou***, Nikhil Gosh*, Bin Yu. "LoRA+: Efficient Low Rank Adaptation of Large Models". In: *International Conference on Machine Learning* (ICML 2024). (*equal contribution).

✧ **Soufiane Hayou**. "$\overleftrightarrow{\text{WD}}$(II): Commutative Width and Depth Scaling in Deep Neural Networks". 2024. In: *Journal of Machine Learning Research* (JMLR).

✧ Greg Yang, Dingli Yu, Chen Zhu, **Soufiane Hayou**. "Tensor Programs VI: Feature Learning in Infinite Depth Neural Networks". In: *International Conference on Learning Representations* (ICLR 2024).

✧ Jiayuan Ye, Anastasia Borovykh, **Soufiane Hayou**, Reza Shokri. "Leave-one-out Distinguishability in Machine Learning". In: *International Conference on Learning Representations* (ICLR 2024).

✧ Mohamed El Amine Seddik, Suei-Wen Chen, **Soufiane Hayou**, Pierre Youssef, Merouane Debbah. "How Bad is Training on Synthetic Data? A Statistical Analysis of Language Model Collapse". In: *Conference on Language Modelling* (COLM 2024).

✧ **Soufiane Hayou**, Greg Yang. "$\overleftrightarrow{\text{WD}}$(I): Width and Depth Limits commute in Residual Networks". In: *International Conference on Machine Learning* (ICML 2023).

✧ Fadhel Ayed*, **Soufiane Hayou***. "Data Pruning and Neural Scaling Laws: Fundamental Limitations of Score-based Algorithms". 2022. In: *Transactions on Machine Learning Research* (TMLR). (*equal contribution, alphabetical order).

✧ **Soufiane Hayou**. "On the Infinite-depth Limit of Finite-width Neural Networks". 2022. In: *Transactions on Machine Learning Research* (TMLR).

✧ Yizhang Lou, Chris Mingard, Yoonsoo Nam, **Soufiane Hayou**. "Feature Learning and Signal Propagation in Deep Neural Networks". In: *International Conference on Machine Learning* (ICML 2022, Spotlight).

✧ Fusheng Liu, Qianxiao Li, **Soufiane Hayou**, Haizhao Yang. "From Optimization Dynamics to Generalization Bounds via Lojasiewicz Gradient Inequality". 2022. In: *Transactions on Machine Learning Research* (TMLR).

✧ **Soufiane Hayou**, Fadhel Ayed. "Regularization in ResNet with Stochastic Depth". In: *Neural Information Processing Systems* (NeurIPS 2021).

✧ **Soufiane Hayou**, Bobby He, Gintare Karolina Dziugaite. "Probabilistic Fine-tuning of Pruning Masks

and PAC-Bayes Self-bounded Learning". (2021, arXiv).

✧ **Soufiane Hayou**, Jean-Francois Ton, Arnaud Doucet, Yee Whye Teh. "Robust Pruning at Initialization". In: *International Conference on Learning Representations* (ICLR 2021).

✧ **Soufiane Hayou***, Eugenio Clerico*, Bobby He*, George Deligiannidis, Arnaud Doucet, Judith Rousseau. "Stable ResNet". In: *The International Conference on Artificial Intelligence and Statistics* (AISTATS 2021, Oral presentation). (*equal contribution).

✧ **Soufiane Hayou**, Arnaud Doucet, Judith Rousseau. "On the Impact of the Activation function on Deep Neural Networks Training". In: *International Conference on Machine Learning* (ICML 2019).

✧ **Soufiane Hayou**, Arnaud Doucet, Judith Rousseau. "Mean-field Behaviour of Neural Tangent Kernel for Deep Neural Networks". (2020, ArXiv).

→ **Workshop papers**:

✧ Greg Yang, Dingli Yu, Chen Zhu, **Soufiane Hayou**. "Feature Learning in Infinite Depth Neural Networks". In: *NeurIPS 2023 Workshop on Mathematics of Modern Machine Learning (M3L)*. Full paper published at ICLR 2024.

✧ Fadhel Ayed*, **Soufiane Hayou***. "Data Pruning and Neural Scaling Laws: Fundamental Limitations of Score-based Algorithms". In: *ICML 2024 Workshop on Data-Centric Machine Learning Research*. (*equal contribution, alphabetical order). Full paper published at Transactions on Machine Learning Research (TMLR).

✧ Fadhel Ayed, **Soufiane Hayou**. "The curse of (non)convexity: The case of an Optimization-Inspired Data Pruning algorithm". In: *NeurIPS 2022 ICBINB workshop*.

✧ **Soufiane Hayou**, Arnaud Doucet, Judith Rousseau. "The Curse of Depth in Kernel Regime". In: *NeurIPS 2021 ICBINB workshop (Spotlight)*. Also published as part of the workshop volume at *Proceedings of Machine Learning Research (PMLR)*.

✧ **Soufiane Hayou**, Bobby He, Gintare Karolina Dziugaite. "Stochastic Pruning: Fine-Tuning, and PAC-Bayes Bound Optimization". In: *NeurIPS 2021 Bayesian Deep Learning Workshop*.

→ **Notes and Technical Reports**:

✧ **Soufiane Hayou**. "On the Connection Between Riemann Hypothesis and a Special Class of Neural Networks". (2023, arXiv).

✧ **Soufiane Hayou**. "On the Overestimation of the Largest Eigenvalue of a Covariance Matrix". (2017, Bloomberg).

✧ **Soufiane Hayou**. "Cleaning the Correlation Matrix with a Denoising AutoEncoder". (2017, Bloomberg).

Teaching Experience

Department of Mathematics, National University of Singapore
Instructor *2021-2023*
○ Principles of Machine Learning (DSA5105, graduate level, 100+ enrolled students).
○ Advanced Topics in Machine Learning (DSA5202, graduate).

Department of Statistics, University of Oxford

Class Tutor / Teaching Assistant   *2018-2020*
- Advanced Simulation Methods (SC5, undergraduate).
- Bayesian Inference and MCMC methods (SC7, graduate).

Awards and Honors

- 2024: Gradient AI Fellowship ($50k in cloud credits, one awarded from over 1000 applications).
- 2023: Google Cloud TPU Credits Award ($10k in cloud credits).
- 2023: Selected as a Rising Star in AI by KAUST.
- 2022: Top Reviewer at NeurIPS.
- 2022: SlowDNN Workshop Travel Grant ($1.5k).
- 2021: Institute of Mathematical Science (NUS) Travel Award ($5k).
- 2019: James Fund for Mathematics Travel Grant (£2k , one of two awarded campus-wide from over 100 applications).
- 2019: ICML 2019 Student Travel Award ($1.5k).
- 2018: Natixis Best Master Thesis in Quantitative Finance (€10k, one of two awarded from over 1000 submissions).
- 2017 - 2021: ESPRC Fellowship Grant ($\sim$ £110k).
- 2017 - 2021: RCUK Fellowship Grant ($\sim$ £42k, awarded by St John's College).
- 2013 - 2017: Egide Grant (€100k, one of twenty awarded by the French Ministry of Foreign Affairs from over 1000 applications).
- 2011: Maroc Telecom Excellence Grant (€20k, awarded to top 1% students in Morocco).
- 2011: Moroccan Government Excellence Scholarship (€10k).

Selected Invited Talks

- August 2024: Nvidia Research, San Francisco.
- May 2024: ThinkAI Hackathon, UM6P, Marrakesh.
- March 2024: AI Seminar, Uber, San Francisco.
- February 2024: AI Seminar, MBZUAI, Abu Dhabi.
- October 2023: MoDL Seminar, Simons Institute, UC Berkeley.
- June 2023: FLAIR Seminar, EPFL, Switzerland.
- May 2023: Morocco AI Webinar (virtual).
- March 2023: Nvidia Research, Hong Kong.
- February 2023: AI Seminar at Technology Innovation Institute, Abu Dhabi.
- January 2023: One World Seminar Series on the Mathematics of Machine Learning (virtual).
- December 2022: Computational and Methodological Statistics (CMStatistics 2022), London.
- November 2022: Rough Paths Seminar, The Alan Turing Institute, London.
- September 2022: 3rd Symp on Machine Learning and Dynamical Systems, Fields Institute, Toronto.
- May 2022: Abu Dhabi Stochastics Seminar, Department of Mathematics, NYU Abu Dhabi.
- Feb 2022: CDSML Seminar Series, National University of Singapore.
- Dec 2021: Seminar Series, Department of Statistics, University of Toronto (virtual).

- Nov 2020: Department of Statistics Seminar, University of Oxford.

Academic Services

- Reviewing for conferences: NeurIPS, ICML, ICLR, AISTATS, UAI.
- Reviewing for journals: Journal of Machine Learning Research (JMLR), Transactions on Machine Learning Research (TMLR), Journal of Computational and Graphical Statistics, Bernoulli.
- Co-organizer: Machine Learning and its Applications (2022, Institute for Mathematical Science (NUS) and Simons Institute for the Theory of Computing), Deep learning reading group (2018-2020, Oxford statistics).

Industry Experience

G-Research   **London, UK**
Quantitative Research Intern   *July 2019 – September 2019*
- Developed deep learning models for the estimation of the correlation matrix and robust portfolio optimization.

Bloomberg LP   **New York, USA**
Quantitative Research Intern   *March 2017 – August 2017*
- Developed scalable ML algorithms for financial data analysis.
- Conducted research on the empirical correlation matrix.
- Developed a Denoising Autoencoder model for the estimation of the correlation matrix.

J.P. Morgan   **London, UK**
Quantitative Research Intern   *April 2016 – September 2016*
- Calibrated a Stochastic model for interest rate products.
- Designed an Initial Margin pricer for interest rate swaps.

Technical Skills

Coding
- Python packages: Utils (Numpy, Scipy, Pandas etc.)
- Deep Learning: PyTorch (6+ years), HuggingFace Transformers (2+ years), LangChain (1+ years), RAG (vectorDB)
- Cloud Platforms: AWS, Google Cloud

Languages
- Moroccan: native
- Arabic: native
- French: bilingual
- English: fluent
- Spanish: beginner