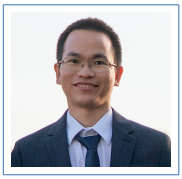
TrungTin Nguyen

Curriculum Vitae (Long)

Postdoc at Statify Team - Inria
✉ *trung-tin.nguyen@inria.fr*
Homepage: *trung-tinnguyen.github.io*
Short CV Version, November 14, 2023



"The book of nature is written in the language of mathematics." (Galileo G., 1890).
"Essentially, all models are wrong, but some models that know when they are wrong, are useful." (adapted from George E.P. Box, 1979).

Academic Appointment

04/12/2023–03/12/2024 **Postdoctoral Research Fellow**, *School of Mathematics and Physics*, *University Of Queensland*, Brisbane, Australia.
Topic: Mathematical analysis of operator learning with artificial neural networks.
Mentors: Hien Duy Nguyen, and Xin Guo.

01/05/2023–30/09/2023 **Postdoctoral Research Fellow (Contrat UGA, MIAI Grenoble Alpes)**, *Statify Team*, *Inria centre at the University Grenoble Alpes*, Grenoble, France.
Topic: Bayesian model selection and simulated-based inference for complex and high-dimensional models.
Mentors: Florence Forbes, and Julyan Arbel.

01/01/2022–30/04/2023 **Postdoctoral Research Fellow (Contrat Inria)**, *Statify Team*, *Inria centre at the University Grenoble Alpes*, Grenoble, France.
+01/11/2023–30/11/2023 **Topic**: Bayesian model selection and simulated-based inference for complex and high-dimensional models.
Mentors: Florence Forbes, and Julyan Arbel.

Education

2018–2021 **Doctor of Philosophy**, *Normandie Université*, Caen, France.
Major in Statistics and Data Science. Defended on December 14, 2021.
Thesis title: Model selection and approximation in high-dimensional mixture of experts models: from theory to practice.
Advisors: Faïcel Chamroukhi. **Rapporteurs**: Sylvain Arlot, and Judith Rousseau.
Committee members: Christophe Biernacki, Hien Duy Nguyen, and Gaëlle Chagny.

2017–2018 **Master of Science, Technology and Health**, *Université d'Orléans*, Orléans, France, GPA: 18/20.
Mention: *Très Bien*. Major in Applied Mathematics.
Thesis title: Reinforcement learning for resource allocation problems using a partially observable Markov decision process.
Advisor: Le Thi Hoai An.

2013–2017 **Bachelor of Science**, *Vietnam National University-Ho Chi Minh Univeristy of Science (VNU-HCM)*, Ho Chi Minh City, Vietnam, GPA: 9.17/10. Rank: 2/1557, Summa Cum Laude.
Honors Program in Mathematics and Computer Science. Major in Probability and Statistics.
Thesis title: Multiplicative censoring model.
Advisor: Dang Duc Trong.

2010–2013 **High School for the Gifted**, *Hung Vuong High School for the Gifted*, Binh Duong, Vietnam, Summa Cum Laude.

2006–2010 **Secondary School**, *Nguyen Quoc Phu Secondary School*, Binh Duong, Vietnam, Summa Cum Laude.

2001–2006 **Primary School**, *Tan Vinh Hiep A Primary School*, Binh Duong, Vietnam, Summa Cum Laude.

Research Interests

Statistical learning Model selection (minimal penalties and slope heuristics, non-asymptotic oracle inequalities), simulation-based inference (approximate Bayesian computation, Bayesian synthetic likelihood, method of moments), Bayesian nonparametrics (Gibbs-type priors, Dirichlet process mixture), high-dimensional statistics (variable selection via Lasso and penalization, graphical models), uncertainty estimation.

Machine learning Supervised learning (deep hierarchical mixture of experts, deep neural networks), unsupervised learning (clustering via mixture models, dimensionality reduction via principal component analysis, deep generative models via variational autoencoders, generative adversarial networks and normalizing flows), reinforcement learning (partially observable Markov decision process).

Optimization Robust and effective optimization algorithms for mixture models (expectation–maximization, variational Bayesian expectation–maximization, Markov chain Monte Carlo methods), difference of convex algorithm, optimal transport (Wasserstein distance, voronoi loss function).

Applications Natural language processing (large language model), remote sensing (planetary science, e.g., retrieval of Mars surface physical properties from hyper-spectral images), audio processing (sound source localization), biostatistics (genomics, transcriptomics, proteomics), computer vision (image segmentation).

Publications

Total 5 Journal Publications + 7 Conference Publications + 5 Preprints.

Deep neural networks

2023 Truong Giang Do, Huy Khiem Le, Quang Pham, TrungTin Nguyen, Binh T. Nguyen, Thanh-Nam Doan, Chenghao Liu, Savitha Ramasamy, Xiaoli Li, and Steven HOI. HyperRouter: Towards Efficient Training and Inference of Sparse Mixture of Experts. In ***Empirical Methods in Natural Language Processing. EMNLP 2023 Main***, December 2023.

Approximation capabilities and convergence rates of the mixture of experts models

2023 Huy Nguyen, TrungTin Nguyen, Khai Nguyen, and Nhat Ho. Towards Convergence Rates for Parameter Estimation in Gaussian-gated Mixture of Experts. *arXiv preprint arXiv:2305.07572*, May 2023.

2023 Huy Nguyen, TrungTin Nguyen, and Nhat Ho. Demystifying Softmax Gating Function in Gaussian Mixture of Experts. In ***Thirty-seventh Conference on Neural Information Processing Systems. NeurIPS 2023 Spotlight***, 2023.

2023 Huy Nguyen, Pedram Akbarian, TrungTin Nguyen, and Nhat Ho. A General Theory for Softmax Gating Multinomial Logistic Mixture of Experts. *arXiv preprint arXiv:2310.14188*, October 2023.

2022 TrungTin Nguyen, Faicel Chamroukhi, Hien D. Nguyen, and Geoffrey J. McLachlan. Approximation of probability density functions via location-scale finite mixtures in Lebesgue spaces. ***Communications in Statistics - Theory and Methods***, pages 1–12, May 2022.

2021 Hien Duy Nguyen, TrungTin Nguyen, Faicel Chamroukhi, and Geoffrey John McLachlan. Approximations of conditional probability density functions in Lebesgue spaces via mixture of experts models. ***Journal of Statistical Distributions and Applications***, volume 8, page 13, 2021.

2020 TrungTin Nguyen, Hien D Nguyen, Faicel Chamroukhi, and Geoffrey J McLachlan. Approximation by finite mixtures of continuous density functions that vanish at infinity. ***Cogent Mathematics & Statistics***, volume 7, page 1750861. Cogent OA, 2020.

Model selection

2023 TrungTin Nguyen, Dung Ngoc Nguyen, Hien Duy Nguyen, and Faicel Chamroukhi. A non-asymptotic risk bound for model selection in high-dimensional mixture of experts via joint rank and variable selection. In ***Australasian Joint Conference on Artificial Intelligence 2023. AJCAI 2023 Long Oral Presentation***, Brisbane, Australia, November 2023.

2022 TrungTin Nguyen, Hien Duy Nguyen, Faicel Chamroukhi, and Florence Forbes. A non-asymptotic approach for model selection via penalization in high-dimensional mixture of experts models. ***Electronic Journal of Statistics***, volume 16, pages 4742 – 4822, 2022. Publisher: Institute of Mathematical Statistics and Bernoulli Society.

2022 TrungTin Nguyen, Faicel Chamroukhi, Hien Duy Nguyen, and Florence Forbes. Model selection by penalization in mixture of experts models with a non-asymptotic approach. In *JDS 2022 - 53èmes Journées de Statistique de la Société Française de Statistique (SFdS)*, Lyon, France, June 2022.

2021 TrungTin Nguyen, Faicel Chamroukhi, Hien Duy Nguyen, and Florence Forbes. Non-asymptotic model selection in block-diagonal mixture of polynomial experts models. *arXiv preprint arXiv:2104.08959.*, 2021.

2021 TrungTin Nguyen. *Model Selection and Approximation in High-dimensional Mixtures of Experts Models: From Theory to Practice*. Ph.D. Thesis, Normandie Université, December 2021.

2020 TrungTin Nguyen, Hien D Nguyen, Faicel Chamroukhi, and Geoffrey J McLachlan. An l_1-oracle inequality for the Lasso in high-dimensional mixtures of experts models. *arXiv preprint arXiv:2009.10622*, 2020.

Bayesian nonparametrics

2023 TrungTin Nguyen, Florence Forbes, Julyan Arbel, and Hien Duy Nguyen. Bayesian nonparametric mixture of experts for high-dimensional inverse problems. *Preprint. hal-04015203*, March 2023.

2022 TrungTin Nguyen, Florence Forbes, and Julyan Arbel. Bayesian nonparametric mixture of experts for high-dimensional inverse problems. In *BNP13 – 13th Conference on Bayesian Nonparametrics*, Puerto Varas, Chile, 2022.

Simulation-based inference

2023 Hien Duy Nguyen, TrungTin Nguyen, Julyan Arbel, and Florence Forbes. Concentration results for approximate Bayesian computation without identifiability. *Preprint. hal-03987197*, February 2023.

2022 Florence Forbes, Hien Duy Nguyen, TrungTin Nguyen, and Julyan Arbel. Supporting Information Summary statistics and discrepancy measures for approximate Bayesian computation via surrogate posteriors. ***Statistics and Computing***, volume 32, page 85, October 2022.

2022 Florence Forbes, Hien Duy Nguyen, TrungTin Nguyen, and Julyan Arbel. Mixture of expert posterior surrogates for approximate Bayesian computation. In *JDS 2022 - 53èmes Journées de Statistique de la Société Française de Statistique (SFdS)*, Lyon, France, June 2022.

2022 Florence Forbes, Hien Duy Nguyen, TrungTin Nguyen, and Julyan Arbel. Summary statistics and discrepancy measures for approximate Bayesian computation via surrogate posteriors. ***Statistics and Computing***, volume 32, page 85, October 2022.

2021 Julyan Arbel, Florence Forbes, Hien Duy Nguyen, and TrungTin Nguyen. Approximate Bayesian computation with surrogate posteriors. In *ISBA 2021 - World Meeting of the International Society for Bayesian Analysis*, Marseille, France, June 2021.

Fellowships and Awards

10/2022 BNP13 Junior Travel Award (1000 USD) granted by International Society for Bayesian Analysis.

2022-2025 Qualifications aux fonctions de Maître de Conférences. Section 26. Mathématiques appliquées et applications des mathématiques. **Designated rapporteurs**: Fabienne Comte, and Fanny Villers.

2022–2023 Postdoctoral Fellowships granted by Inria centre at the University Grenoble Alpes, France.

2018–2021 Ph.D. Scholarship granted by Ministère de l'Enseignement Supérieur et de la Recherche, France.

2017 Highest Distinction Graduation Award and Outstanding Student Award, VNU-HCM, Vietnam

2014–2017 Scholarship of the National Program for the Development of Mathematics 2010–2020 of Vietnam Institute for Advanced Study in Mathematics (VIASM), Vietnam.

Teaching Experiences

01-04/2023: **Statistical analysis and document mining (Complementary Course, 16.5h). Responsible professor: Pedro Rodrigues**, *Master 1 of Applied Mathematics, Université Grenoble Alpes*, France.

09-12/2022: **Méthodes statistiques pour la biologie - STA301 (Travaux Dirigés, 23h). Responsible professor: Julien Chevallier**, *Licence Sciences et Technologies - BIO, Université Grenoble Alpes*, France.

Fall 2018: **Mathematical and numerical foundations of modeling and simulation using partial differential equations (Teaching Assistant, 24h). Responsible professor: Jing-Rebecca Li (IDEFIX team, Inria)**, *French-Vietnam Master 2 in Applied Mathematics, VNU-HCM*, Vietnam.

Fall 2017: **Principles of Mathematical Analysis (Teaching Assistant, 30h). Responsible professor: Duong Minh Duc**, *Bachelor in Mathematics and Computer Science, VNU-HCM*, Vietnam.

Collaborators (in random order)

1. Geoff McLachlan: Professor of Statistics (Personal Chair), School of Mathematics and Physics, University Of Queensland, Brisbane, Australia.
2. Florence Forbes: Senior Researcher (Director of Research), Statify Team, Inria centre at the University Grenoble Alpes, Grenoble, France.
3. Faïcel Chamroukhi: Full Professor of Statistics and Data Science, Université de Caen Normandie, Lab. of Mathematics Nicolas Oresme (LMNO), UMR CNRS 6139; Scientific head of Data Science and Aritificial Intelligence, IRT SystemX, the Research & Technology Organisation of Université Paris-Saclay, Palaiseau, France.
4. Hien Duy Nguyen: Associate Professor, School of Computing, Engineering and Mathematical Sciences, La Trobe University, Bundoora, Victoria, Australia.
5. Nhat Ho: Assistant Professor, Department of Statistics and Data Sciences, The University of Texas at Austin, Austin, Taxas, United States.
6. Julyan Arbel: Associate Researcher (Chargé de Recherche), Statify Team, Inria centre at the University Grenoble Alpes; Laboratoire Jean Kuntzmann, member of Université Grenoble Alpes Grenoble, Grenoble, France.
7. Binh Nguyen: Associate Professor of Computer Science and the Head of the Department of Computer Science, Faculty of Mathematics and Computer Science, University of Science (VNUHCM-US), Vietnam National University (VNUHCM), Ho Chi Minh City, Vietnam.
8. Quang Pham: Research Scientist, Machine Intellection department, Institute for Infocomm Research (I2R), A*Star, Singapore.
9. Huy Nguyen: Phd Student, Department of Statistics and Data Sciences, The University of Texas at Austin, Austin, Taxas, United States.
10. Khai Nguyen: Phd Student, Department of Statistics and Data Sciences, The University of Texas at Austin, Austin, Taxas, United States.
11. Dung Ngoc Nguyen: Postdoctoral Research Fellow in Statistics, Department of Statistical Sciences, University of Padova, Padova, Italy.
12. Ho Minh Duy Nguyen: PhD Candidate, Max Planck Research School for Intelligent Systems & DFKI, Stuttgart, Germany.
13. Giang Truong Do: Research Assistant, University of Tennesse, Tennessee, United States.
14. Le Huy Khiem: Research Assistant, Vin University, Hanoi, Vietnam.

Conference, Seminar, Workshop Presentations

08/2023 Summary statistics and discrepancy measures for approximate Bayesian computation via surrogate posteriors at The 10th Vietnam Mathematical Congress, Da Nang, Vietnam (Oral presentation).

07/2023 Summary statistics and discrepancy measures for approximate Bayesian computation via surrogate posteriors at Summer school on Bayesian statistics and computation, Ho Chi Minh, Vietnam (Poster presentation).

12/2022 A non-asymptotic approach for model selection via penalization in high-dimensional mixture of experts models at 2022 IMS International Conference on Statistics and Data Science (ICSDS), Florence, Italy (Oral presentation).

11/2022 A non-asymptotic approach for model selection via penalization in high-dimensional mixture of experts models at Séminaire Données et Aléatoire Théorie & Applications, Laboratoire Jean Kuntzmann, Grenoble, France (Invited Speaker).

10/2022 Bayesian nonparametric mixture of experts for high-dimensional inverse problems at BNP13 – 13th Conference on Bayesian Nonparametrics, Puerto Varas, Chile (Oral presentation).

06/2022 Model selection by penalization in mixture of experts models with a non-asymptotic approach at JDS 2022 - 53èmes Journées de Statistique de la Société Française de Statistique (SFdS), Lyon, France (Oral presentation).

05/2022 A non-asymptotic approach for model selection via penalization in high-dimensional mixture of experts models at Seminar on Applied Statistics, Vietnam Institute for Advanced Study in Mathematics, Vietnam (Invited Speaker).

04/2022 A non-asymptotic approach for model selection via penalization in mixture of experts models at Statlearn 2022, Institut d'Etudes Scientifiques de Cargèse, Corsica (Poster presentation)

03/2022 A non-asymptotic model selection in mixture of experts models at Séminaire de Statistique Rennais, ENSAI École Nationale de Statistique et Analyse de l'Information, Rennes, France (Invited Speaker).

10/2021 Model Selection and Approximation in High-dimensional mixture of experts Models: From Theory to Practice at Jed 2021: Journée scientifique de l'École Doctorale 2021, Le Havre, France (Oral presentation).

09/2021 Approximation and non-asymptotic model selection in mixture of experts models at Journée Thématique: "Intelligence Artificielle - Applications et défis mathématiques", INSA Rouen Normandie, Rouen, France (Poster session).

06/2021 Non-asymptotic model selection in mixture of polynomial experts models at MHC2021 Mixtures Hidden Markov model Clustering, Institut de Mathématique d'Orsay, Paris, France (Poster session).

04/2021 Non-asymptotic model selection for the Gaussian-gated localized mixture of experts regression models at MiMo 2021: Workshop on Mixture Models, Laboratoire de Mathématiques Raphaël Salem, Université de Rouen Normandie, France (Invited speaker).

Professional Services

Journal Reviewing

Journal of the American Statistical Association (Taylor Francis): 1 paper. See certificate.

Australian & New Zealand Journal of Statistics (Wiley): 2 paper. See certificate.

Computational Statistics and Data Analysis (Elsevier): 4 papers. See certificate.

Neurocomputing (Elsevier): 1 paper. See certificate.

Biometrical Journal (Wiley): 2 papers. See certificate.

Communications in Statistics - Theory and Methods (Taylor Francis): 2 papers. See certificate.

Conference Reviewing/Program Committee

Proceedings of the Research School on Statistics and Data Science (RSSDS 2019) (Springer): 2 papers

Editorial Board

International Journal of Machine Intelligence and Sensory Signal Processing (Inderscience): Associate Editors

Projects

2023–2026 **Member of the WOMBAT (Variance-reduced Optimization Methods and Bayesian Approximation Techniques for scalable inference)**.

- Principal investigator:
  - Florence Forbes (Statify Inria Grenoble, France),
  - Hien Duy Nguyen (School of Computing, Engineering and Mathematical Sciences, La Trobe University, Bundoora, Victoria, Australia).
- Other participants:
  - Queensland University of Technology, Brisbane, Australia,
  - University of Queensland, Brisbane, Australia,
  - Swinburne University of Technology, Melbourne, Australia
  - The University of Adelaide, Australia.
  - Université de Caen Normandie, France.
- Website: `https://team.inria.fr/statify/projects/WOMBAT/`

2019–2021 **Member of the LANDER (Latent Analysis, Adversarial Networks, and DimEnsionality Reduction)**.

- Principal investigator:
  - Florence Forbes (Mistis Inria Grenoble Rhone-Alpes, France),
  - Hien Duy Nguyen (School of Mathematics and Physics, University of Queensland, Australia).
- Other participants:
  - Queensland University of Technology, Brisbane, Australia,
  - Swinburne University of Technology, Melbourne, Australia
  - Université de Caen Normandie, France.
- Website: `https://team.inria.fr/statify/projects/lander/`
- My contributions in this project: A non-asymptotic approach for model selection via penalization in high-dimensional mixture of experts models. Electronic Journal of Statistics, 2022.

Selected Academic Experiences

12/2022 **Accomplished an online course Machine Learning Specialization**, *Stanford University, USA*, instructed by Professor Andrew Ng et al..

Including 3 courses:

- Supervised Machine Learning: Regression and Classification. Grade: 100%.
- Advanced Learning Algorithms. Grade: 100%.
- Unsupervised Learning, Recommenders, Reinforcement Learning. Grade: 100%.

Course Certificates:
`https://coursera.org/share/a9473e1b59c38bbde2f413bed53f3ebf`

07/2019 **Participated in 3rd International Summer School on Deep Learning (39 hours)**, *Warsaw, Poland*.

Including some featured courses:

- *Deep Generative Models* by Aaron Courville (University of Montréal, Canada).
- *Dive into Deep Learning* by Alex Smola (Amazon, USA).
- *Mathematics of Deep Learning* by Rene Vidal (Johns Hopkins University, USA).

06-09/2018 **Accomplished an online course Deep Learning Specialization**, *Stanford University, USA*, instructed by Professor Andrew Ng et al..

Including 5 courses:

- Neural Networks and Deep Learning. Grade: 100%.
- Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization. Grade: 100%.
- Structuring Machine Learning Projects. Grade: 98.3%.
- Convolutional Neural Networks. Grade: 98.9%.
- Sequence Models. Grade: 100%.

Course Certificates:
`https://coursera.org/share/3d5d2ffa4a112d75883b62a22d4a132c`

Professional Memberships

08/2021 Institute of Mathematical Statistics (IMS).

01/2020 International Society for Bayesian Analysis (ISBA).

01/2022 Société Française de Statistique (SFdS).

Languages

Vietnamese **Mother tongue**

English **IELTS 7.0/9.0** *Excellent reading and listening skills, good at writing and speaking.*

French **Intermediate B1-B2** *Good reading and writing skills, can understand isolated sentences and common phrases in listening and speaking.*

Computer Skills

Programming Languages: Advanced R, Advanced Python, MATLAB, C++, SAS.

Operating Systems: Linux, macOS, Microsoft Windows.

Referees

Florence Forbes
Senior Researcher
Head of the Statify team
Inria centre at the University Grenoble Alpes, France
✉ florence.forbes@inria.fr

Geoffrey McLachlan
Professor of Statistics
School of Mathematics and Physics
University of Queensland, Australia
✉ g.mclachlan@uq.edu.au

Hien Duy Nguyen
Associate Professor
School of Computing, Engineering and Mathematical Sciences
La Trobe University, Australia
✉ h.nguyen5@latrobe.edu.au

Nhat Ho
Assistant Professor
Department of Statistics and Data Sciences
The University of Texas at Austin, United States
✉ minhnhat@utexas.edu

Faïcel Chamroukhi
Professor of Statistics and Data Science
Head of Data Science and Artificial Intelligence
IRT SystemX, France
✉ Faicel.chamroukhi@irt-systemx.fr

Julyan Arbel
Associate Researcher
Associate Researcher of the Statify team
Inria centre at the University Grenoble Alpes, France
✉ julyan.arbel@inria.fr

Le Thi Hoai An
Full Professor of Exceptional Class
Director of Computer science and Applications Department
University of Lorraine, France
✉ hoai-an.le-thi@univ-lorraine.fr