NGHIA HOANG

Ph.D. in Computer Science
Senior Machine Learning Scientist,
AWS AI Labs, Amazon
E-mail: tnhoang@amazon.com
Personal E-mail: hoangtrongnghia87@gmail.com
Homepage: https://htnghia87.github.io/
My Google Scholar Profile:
https://scholar.google.com.sg/citations?user=E-kZZeQAAAAJ&hl=en

Research Interests

1. Model-Centric Collaborative Machine Learning, Meta Learning, Personalized Federated Learning

Sample publications: see [S5] **AAAI-19,** [S4] **ICML-19,** [C7] **NeurIPS-19,** [S1] **ICML-20 &** [S2] **ICML-21**.

2. Applications in ML for Healthcare

Sample publications: see [C6] **AAAI-20,** [J1] **TKDE-20 &** [C2] **WWW-21.**
Example work mentioned on media: Drug-Drug Interaction (DDI):

https://www.technologyreview.com/f/615153/ai-adverse-drug-interactions-chemistry-health-care/

3. Bayesian Optimization and Cost-Effective Learning of Gaussian Processes

Sample publications:
see [S8] **ICML-14,** [S7] **ICML-15,** [S6] **ICML-16,** [C19] **AAAI-16,** [C17] **AAAI-17,** [C16] **AAAI-18 &** [C4] **NeurIPS-20.**

Funded Research Proposals

[1] On-Device Personalization with Meta Learning
(IBM PI) [with Patrick Jaillet (MIT PI)]

The research funding (**$150K**) was reviewed and approved by the MIT-IBM board of directors under the exploratory challenge theme of the MIT Quest for Intelligence.

[2] Cross-Species Translation of COVID-19 Systems Serology Data for Infection and Vaccine Treatment
(IBM Co-PI) [with Douglas A. Lauffenburger (MIT PI) and Sara Magliacane (IBM PI)]

The research funding (**$150K**) was reviewed and approved by the MIT-IBM board of directors under the exploratory challenge theme of the MIT Quest for Intelligence.

[3] Toward Trustable Model-centric Sharing for Collaborative Machine Learning
(Affiliated Collaborator) [with Patrick Jaillet (oversea PI)]

The research funding (**S$8.4M**) was approved under the AI Singapore Research Programme.

Education Background

Doctor of Philosophy (Ph.D. in Computer Science) **[2010 – 2015]**

Department of Computer Science, School of Computing (SoC), National University of Singapore (NUS)
Thesis Title: New Advances on Bayesian and Decision-Theoretic Approaches for Interactive Machine Learning.

Following my PhD, I spent the next two years working as a Research Fellow at NUS, and then Postdoctoral Research Associate at MIT. After my postdoctoral training, I joined MIT-IBM Watson AI Lab as a Research Staff Member in 2018. In 2021, I joined the Amazon AWS AI Labs as a Senior Machine Learning Scientist.

Selected Publications

[S1] Chi Thanh Lam, Trong Nghia Hoang, Kian Hsiang Low and Patrick Jaillet (2021), **"Model Fusion for Personalized Learning"**. In the Proceeding of the International Conference on Machine Learning **(ICML-21)**. **(co-first author)** [21.47% acceptance rate]

[S2] Trong Nghia Hoang, Chi Thanh Lam, Kian Hsiang Low and Patrick Jaillet (2020), **"Learning Task-Agnostic Embedding of Multiple Black-Box Experts for Multi-Task Model Fusion"**. In the Proceeding of the International Conference on Machine Learning **(ICML-20)**. [21.08% acceptance rate]

[S3] Minh Hoang, Trong Nghia Hoang, Hai Pham and David Woodruff (2020), **"Revisiting the Sample Complexity of Sparse Spectrum Approximation of Gaussian Processes".** In the Proceeding of Advances in Neural Information Processing Systems **(NeurIPS-20)**. **(co-first author)** [20.06% acceptance rate]

[S4] Mikhail Yurochkin, Mayank Argawal, Soumya Ghosh, Kristjan Greenewald, Trong Nghia Hoang and Yasaman Khazaeni (2019), **"Bayesian Nonparameteric Federated Learning for Neural Networks"**. In the Proceeding of the International Conference on Machine Learning **(ICML-19)**. [22.60% acceptance rate]

[S5] Trong Nghia Hoang, Quang Minh Hoang, Kian Hsiang Low and Jonathan How (2019), **"Collective Online Learning via Decentralized Gaussian Processes in Massive Multi-Agent Systems".** In the Proceeding of the 33rd AAAI Conference in Artificial Intelligence **(AAAI-19)**. [16.20% acceptance rate]

[S6] Trong Nghia Hoang, Quang Minh Hoang and Kian Hsiang Low (2016), **"A Distributed Variational Inference Framework for Unifying Parallel Sparse Gaussian Process Regression Models".** In the Proceeding of the 33rd International Conference on Machine Learning **(ICML -16).** [24.30% acceptance rate]

[S7] Trong Nghia Hoang, Quang Minh Hoang and Kian Hsiang Low (2015), **"A Unifying Framework of Anytime Sparse Gaussian Process Regression Models with Stochastic Variational Inference for Big Data"**. In the Proceeding of the 32nd International Conference on Machine Learning **(ICML -15).** [26.0% acceptance rate]

[S8] Trong Nghia Hoang, Kian Hsiang Low, Patrick Jaillet and Mohan Kankanhalli (2014), **"Nonmyopic ϵ-Bayes-Optimal Active Learning of Gaussian Processes"**. In the Proceeding of the 31st International Conference on Machine Learning **(ICML -14).** [Cycle 2, 25.0% acceptance rate]

Conference Publications (Full List)

[C1] Chi Thanh Lam, Trong Nghia Hoang, Kian Hsiang Low and Patrick Jaillet (2021), **"Model Fusion for Personalized Learning"**. In the Proceeding of the International Conference on Machine Learning **(ICML-21)**. **(co-first author)** [21.47% acceptance rate]

[C2] Trong Nghia Hoang, Shenda Hong, Cao Xiao, Bryan Kian Hsiang Low and Jimeng Sun (2021), **"AID: Active Distillation Machine to Leverage Pre-Trained Black-Box Models in Private Data Settings".** In the Proceeding of the Web Conference **(WWW-21)**. [20.60% acceptance rate]

[C3] Nathan Hunt, Nathan Fulton, Sara Magliacane, Trong Nghia Hoang, Subhro Das and Armando Solar-Lezama (2021), **"Verifiably Safe Exploration for End-to-End Reinforcement Learning"**. In the Proceeding of the ACM International Conference on Hybrid Systems: Computation and Control **(HSCC-21).**

[C4] Minh Hoang, Trong Nghia Hoang, Hai Pham and David Woodruff (2020), **"Revisiting the Sample Complexity of Sparse Spectrum Approximation of Gaussian Processes".** In the Proceeding of Advances in Neural Information Processing Systems **(NeurIPS-20)**. **(co-first author)** [20.06% acceptance rate]

[C5] Trong Nghia Hoang, Chi Thanh Lam, Kian Hsiang Low and Patrick Jaillet (2020), **"Learning Task-Agnostic Embedding of Multiple Black-Box Experts for Multi-Task Model Fusion"**. In the Proceeding of the International Conference on Machine Learning **(ICML-20)**. [21.08% acceptance rate]

[C6] Kexin Huang, Cao Xiao, Trong Nghia Hoang, Lucas Glass and Jimeng Sun (2020), **"CASTER: Predicting Drug Interactions with Chemical Substructure Representation"**. In the Proceeding of the AAAI Conference in Artificial Intelligence **(AAAI-20)**. [20.4% acceptance rate]

[C7] Mikhail Yurochkin, Mayank Argawal, Soumya Ghosh, Kristjan Greenewald and Trong Nghia Hoang (2019), **"Statistical Model Aggregation via Parameter Matching"**. In the Proceeding of Advances in Neural Information Processing Systems **(NeurIPS-19)**. [21% acceptance rate]

[C8] Pu Zhao, Sijia Liu, Pin-Yu Chen, Trong Nghia Hoang, Kaidi Xu, Bhavya Kailkhura and Xue Lin (2019), **"On the Design of Adversarial Examples by Leveraging Gradient-Free Optimization and Operator Splitting Method"**. In the Proceeding of the International Conference on Computer Vision **(ICCV-19)**. [25% acceptance rate]

[C9] Quang Minh Hoang, Trong Nghia Hoang, Kian Hsiang Low and Carleton Kingsford (2019), **"Collective Model Fusion for Multiple Black-Box Experts"**. In the Proceeding of the International Conference on Machine Learning **(ICML-19)**. [22.60% acceptance rate] **(co-first author)**

[C10] Mikhail Yurochkin, Mayank Argawal, Soumya Ghosh, Kristjan Greenewald, Trong Nghia Hoang and Yasaman Khazaeni (2019), **"Bayesian Nonparameteric Federated Learning for Neural Networks"**. In the Proceeding of the International Conference on Machine Learning **(ICML-19)**. [22.60% acceptance rate]

[C11] Hong Shenda, Cao Xiao, Trong Nghia Hoang, Tengfei Ma, Hongyan Li and Jimeng Sun (2019), **"RDPD: Rich Data Helps Poor Data via Imitation"**. In the Proceeding of the International Joint Conference on Artificial Intelligence **(IJCAI-19)**. [17.88% acceptance rate]

[C12] Tianfan Fu, Trong Nghia Hoang, Cao Xiao and Jimeng Sun (2019), **"DDL: Deep Dictionary Learning for Predictive Phenotyping"**. In the Proceeding of the International Joint Conference on Artificial Intelligence **(IJCAI-19)**. [17.88% acceptance rate] **(co-first author)**

[C13] Haibin Yu, Trong Nghia Hoang, Kian Hsiang Low and Patrick Jaillet (2019), **"Stochastic Variational Inference for Bayesian Sparse Gaussian Process Regression"**. In the Proceeding of the International Joint Conference on Neural Networks **(IJCNN-19)**.

[C14] Trong Nghia Hoang, Quang Minh Hoang, Kian Hsiang Low and Jonathan How (2019), **"Collective Online Learning via Decentralized Gaussian Processes in Massive Multi-Agent Systems".** In the Proceeding of the 33rd AAAI Conference in Artificial Intelligence **(AAAI-19)**. [16.20% acceptance rate]

[C15] Trong Nghia Hoang, Yuchen Xiao, Kavinayan Sivakumar, Christopher Amato and Jonathan How (2018), **"Near-Optimal Adversarial Policy Switching for Decentralized Asynchronous Multi-Agent Systems".** In the Proceeding of the International Conference on Robotics and Automation **(ICRA-18)**. [40.6% acceptance rate]

[C16] Trong Nghia Hoang, Quang Minh Hoang, Kian Hsiang Low and Ruofei Ouyang (2018), **"Decentralized High-Dimensional Gaussian Process Optimization with Factor Graphs".** In the Proceeding of the 32nd AAAI Conference in Artificial Intelligence **(AAAI-18)**. [24.55% acceptance rate]

[C17] Quang Minh Hoang, Trong Nghia Hoang and Kian Hsiang Low (2017), **"A Generalized Stochastic Variational Bayesian Hyperparameter Learning Framework for Sparse Spectrum Gaussian Process Regression".** In the Proceeding of the 31st AAAI Conference in Artificial Intelligence **(AAAI-17).** [24.6% acceptance rate]

[C18] Trong Nghia Hoang, Quang Minh Hoang and Kian Hsiang Low (2016), **"A Distributed Variational Inference Framework for Unifying Parallel Sparse Gaussian Process Regression Models".** In the Proceeding of the 33rd International Conference on Machine Learning **(ICML -16).** [24.30% acceptance rate]

[C19] Zhang Yehong, Trong Nghia Hoang, Kian Hsiang Low and Mohan Kankanhalli (2016), **"Near-Optimal Active Learning of Multi-Output Gaussian Processes".** In the Proceeding of the 30th AAAI Conference in Artificial Intelligence **(AAAI -16).** [Oral Presentation] [25.75% acceptance rate]

[C20] Trong Nghia Hoang, Quang Minh Hoang and Kian Hsiang Low (2015), **"A Unifying Framework of Anytime Sparse Gaussian Process Regression Models with Stochastic Variational Inference for Big Data"**. In the Proceeding of the 32nd International Conference on Machine Learning **(ICML -15).** [26.0% acceptance rate]

[C21] Trong Nghia Hoang, Kian Hsiang Low, Patrick Jaillet and Mohan Kankanhalli (2014), **"Nonmyopic ϵ-Bayes-Optimal Active Learning of Gaussian Processes"**. In the Proceeding of the 31st International Conference on Machine Learning **(ICML -14).** [Cycle 2, 25.0% acceptance rate]

[C22] Kian Hsiang Low, Jie Chen, Trong Nghia Hoang, Nuo Xu and Patrick Jaillet (2014), **"Recent Advances in Scaling up Gaussian Process Predictive Model for Large Spatiotemporal Data"**. In the Proceeding of the Dynamic Data-driven Environmental Systems Science Conference **(DyDESS -14)**.

[C23] Prabhu Natarajan, Trong Nghia Hoang, Yongkang Wong, Kian Hsiang Low and Mohan Kankanhalli (2014), **"Scalable Decision-Theoretic Coordination and Control for Real-time Active Multi-Camera Surveillance"**. In the Proceeding of the 8th ACM/IEEE International Conference on Distributed Smart Cameras **(ICDSC-14)**.

[C24] Trong Nghia Hoang and Kian Hsiang Low (2013), **"Towards Practical Planning to Predict and Exploit Intentions for Interacting with Self-Interested Agents"**. In the Proceeding of the 23rd International Joint Conference on Artificial Intelligence **(IJCAI -13).** [Oral Presentation] [13.2% acceptance rate]

[C25] Trong Nghia Hoang and Kian Hsiang Low (2013), **"A General Framework for Interacting Bayes-Optimally with Self-Interested Agents using Arbitrary Parametric Model and Model Prior"**. In the Proceeding of the 23rd International Joint Conference on Artificial Intelligence **(IJCAI -13).** [28.0% acceptance rate]

[C26] Prabhu Natarajan, Trong Nghia Hoang, Kian Hsiang Low and Mohan Kankanhalli (2012), **"Decision-Theoretic Coordination and Control for Active Multi-Camera Surveillance in Uncertain, Partially Observable Environments"**. In the Proceeding of the 6th ACM/IEEE International Conference on Distribute Smart Cameras **(ICDSC -12)**.

[C27] Prabhu Natarajan, Trong Nghia Hoang, Kian Hsiang Low and Mohan Kankanhalli (2012), **"Decision-Theoretic Approach for Maximizing Observation of Multiple Targets in Multi-Camera Surveillance"**. In the Proceeding of the Eleventh International Conference on Autonomous Agents and Multi-agent Systems **(AAMAS-12).** [Full Paper, Oral Presentation] [20.4% acceptance rate]

Workshop Publications

[W1] Nathan Fulton, Nathan Hunt, Trong Nghia Hoang and Subro Das (2020), **"Formal Verification of End-to-End Learning in Cyber-Physical Systems: Progress and Challenges"**. In the Proceeding of the NeurIPS-19 Workshop on Safety and Robustness in Decision Making.

[W2] Yehong Zhang, Trong Nghia Hoang, Kian Hsiang Low and Mohan Kankanhalli (2017), **"Information-Based Multi-Fidelity Bayesian Optimization"**. In the Proceeding of the NIPS-17 Workshop on Bayesian Optimization (BayesOpt-17).

[W3] Trong Nghia Hoang, Kian Hsiang Low, Patrick Jaillet and Mohan Kankanhalli (2014), **"Active Learning is Planning: Nonmyopic ∈-Bayes-Optimal Active Learning of Gaussian Processes"**. In T. Calders, F. Esposito, E. Hullermeier, R. Meo, editors, Proceedings of the 7th European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD-14) Nectar (New Scientific and Technical Advances in Research) Track, Part III, LNCS 8726, Springer.

[W4] Trong Nghia Hoang and Kian Hsiang Low (2012), **"Intention-Aware Planning Under Uncertainty for Interacting with Self-Interested, Boundedly Rational Agents"**. In the Proceeding of the Eleventh International Conference on Autonomous Agents and Multi-agent Systems **(AAMAS-12)**. [Extended Abstract]

Journal Publications

[J1] Cao Xiao, Trong Nghia Hoang, Shenda Hong, Tengfei Ma and Jimeng Sun (2020), **"CHEER: Rich Model Helps Poor Models via Knowledge Infusion"**. In the Proceeding of the IEEE Transaction of Knowledge and Data Engineering **(TKDE-20). (co-first author)**

[J2] Hugo Uvegi, Zach Jensen, Trong Nghia Hoang, Brian Traynor, Tunahan Aytas, Richard T. Goodwin, Elsa Olivetti (2021), **"Literature Mining for Alternative Cementitious Precursors and Dissolution Rate Modeling of Glassy Phases"**. In the Proceeding of the Journal of the American Ceramic Society **(JACerS-21).**

Preprints

[P1] Thanh Vinh Vo, Pengfei Wei, Trong Nghia Hoang and Tze-Yun Leong (2021), **"Adaptive Multi-Source Causal Inference"**. CoRR abs/2105.14877 (2021)

[P2] Thanh Vinh Vo, Trong Nghia Hoang, Young Lee and Tze-Yun Leong (2021), **"Federated Estimation of Causal Effects from Observational Data"**. CoRR abs/2106.00456 (2021)

Working Experience

Senior Machine Learning Scientist (Amazon)	Nov 2020 – present
Research Staff Member (MIT-IBM Watson AI Lab, IBM Research)	Aug 2018 – Oct 2020
Postdoctoral Associate (Laboratory for Information and Decision Systems, MIT)	Apr 2017 – Feb 2018
Research Fellow (SeSaMe Centre, Interactive Digital Media Institute, NUS)	Mar 2015 – Mar 2017

Teaching

Teaching Assistant (NUS): CS4248 Natural Language Processing (Semester 1, 2010-2011), CS4246 AI Planning and Decision Making (Semester 1, 2011 - 2012), CS3243 Introduction to AI (Semester 2, 2012 - 2013).

Guest Lecturer: (1) NUS - CS4246 AI Planning and Decision Making (Semester 1, 2015 - 2016); **(2) Harvard** - Lecture on Gaussian Processes and Sparse Approximations [Host: Professor Natesh Pillai]

Gaussian Processes: Theory & Applications - AI Summer School in Singapore (2020)

Talk

Cost Effective and Scalable Non-parametric Bayesian Machine Learning

At: Nanyang Technological University (Singapore) (Oct 2016); Microsoft Research Cambridge (UK) (Feb, 2017); University of Edinburgh (UK) (Mar, 2017); Singapore Management University (Singapore) (Nov, 2017); and IBM Thomas J. Watson Research Center (US) (Oct, 2017)

Collective Online Learning of Gaussian Process in Large Multi-Agent System

At: University of Manchester (UK) (Jan, 2018); and ISE Department, Rutgers University (US) (Oct, 2018)

Active Interpretation Machine for Explainable Healthcare Analytics

At: ISE Department, Rutgers University (US) (Oct, 2019)

Gaussian Processes: Theory & Applications

At: AI Summer School in Singapore (2020)

US Patents

[1] A Method for Combining Pre-Trained Neural Networks into a Memory and Computation Efficient Global Model (Application Number: 16/576927)

[2] System and Method for Model-Free Safe Symbolic Reinforcement Learning from Visual Inputs (Application Number: 16/709633)

Professional Services

Program Committee Member: IROS'12, IJCAI'15, NIPS'16, ICRA'16, ICRA'18, AAMAS '18, IJCAI'18, PAKDD'18, NIPS'18, AAAI'19, ICLR'19, AAMAS'19, ICML'19, NeuIPS'19, IJCAI'19, MRS'19, IJCAI'20, ICRA'20, AAAI'20, ICML'20, ICLR'20, CoRL'20, AAMAS'20, NeurIPS'20, AAAI'21, IJCAI'21, ICLR'21, ICML'21, NeurIPS'21, CoRL'21, ICLR'22

Senior Program Committee Member: IJCAI'21

Journal Reviewer: IEEE-RA-L (2018), Journal of Selected Topics in Signal Processing (2014-2015), T-PAMI (2019), JAIR (2021)

Research Proposal Reviewer: AI Singapore Research Programme (2018-2019)

Workshop Organizer: Practical Bayesian Methods for Big Data (2019) as part of IBM AI Research Week (2019)

https://deeplybayesian.github.io/