

CURRICULUM VITAE

Canh Hao Nguyen (PhD)

POSITION: Senior Lecturer
ADDRESS: Bioinformatics Center, ICR, Kyoto University, Japan
HOMEPAGE: <http://www.bic.kyoto-u.ac.jp/pathway/canhhao/>
EMAIL: canhhao@kuicr.kyoto-u.ac.jp, canhhao@gmail.com

SUMMARY

With breadth and depth in machine learning (ML), I have worked on **novel ML models, methods on graphs** and **applications** in life sciences. Key contributions: a sound method for *semi-supervised learning* on large graphs, a general framework for *learning on hypergraphs*, and clarifying *convex clustering* solutions.

RESEARCH ACHIEVEMENT

I have published on these topics in venues including IJCAI, TPAMI, AISTATS:

1. **Sparse learning**: Elucidating **convex clustering** solution (proving convexity of solutions), a general framework for **hypergraph learning** (generalizing all previous models, enabling sparsity), new models for sparse **multitask learning**.
2. **Learning on graphs**: new methods for **node embeddings** (combining labeled and unlabeled data for semi-supervised learning, using **graph neural networks** (GNNs), using kernels for general graphs), **link prediction** methods (with latent feature interaction encoded in kernels and GNNs), **structured input/output** models, a *provably sound* method for **semi-supervised learning** on large graphs (for global information loss problem on large graphs).
3. **Kernels**: kernel matrix evaluation, **kernels for structured data** (for graphs, on graphs), **feature selection/metric learning** for **Weisfeiler-Lehman kernels**.
4. **Applications**: **dry eye** detection from microscopic videos, **drug side effect** prediction from chemical compounds and protein networks, **molecular identification** from mass spectrometry data and molecular fingerprints, gene/protein **network analyses**, **microbiology**.

WORK EXPERIENCE

<i>Dec 2020 - Dec 2021</i>	Rohto Pharmaceutical Co., Ltd. AI Research Advisor (casual, still joining meetings)
<i>Nov 2019 - now</i>	Bioinformatics Center, ICR, Kyoto University Senior Lecturer
<i>Jun 2013 - Oct 2019</i>	Assistant Professor
<i>Apr 2009 - May 2013</i>	Postdoctoral researcher
<i>Summer 2007</i>	Xerox Research Center Europe, Meylan, France Summer Intern

SKILLS

Java, Python, Matlab, C/C++, scripting, markup, etc.

EDUCATION

Mar 2009	PhD at Japan Advanced Institute of Science and Technology (JAIST) Thesis: "Kernel Design, Evaluation and Learning for Kernel Methods"
Mar 2006	MS at JAIST with thesis: "Imbalanced Data Learning"
Jul 2002	BS in Computer Science at The University of New South Wales, Australia <i>Honours first class</i> with thesis: "Robocup Vision"

LANGUAGES

ENGLISH	Advanced
JAPANESE	N2 (business level)
VIETNAMESE	Native

HONORS AND AWARDS

Nov 2009 - Oct 2011	JSPS Postdoctoral Fellowship for Foreign Researchers
Mar 2009	President's PhD dissertation award, JAIST
Oct 2003 - Mar 2009	Japanese Government scholarship for Master and PhD in Japan
Jul 2002	Honours First class in Computer Science at UNSW (<i>completed a four-year program within three-year time</i>)
Jul 1999 - Aug 2002	Australian Development Scholarship for undergraduate in Australia
Jul 1997	Silver medal at the <i>38th International Mathematical Olympiad</i> in Mar Del Plata, Argentina (high school)
May 1997	First prize at <i>Vietnam National Mathematical Olympiad</i> (high school)

RESEARCH GRANTS

-
- Principal Investigator (PI): Kakenhi Grand-in Aid for Scientific Research C 2018-2021 (4,160,000 yen) and 2022-2025 (4,160,000 yen).
 - Co-PI: Kakenhi Grand-in Aid for Scientific Research B (17,940,000 yen). ONR grant 2019-2022 (99,000 usd) and AOARD grant 2017-2019 (99,000 usd),

SELECTED PUBLICATIONS

-
1. C.H. Nguyen, H. Mamitsuka, "On Convex Clustering: Convexity, Bounding Balls and Characteristics", *Machine Learning*, in press.
 2. P. Petschner, A.D. Nguyen, C.H. Nguyen, H. Mamitsuka, "Machine learning for predicting drug-drug interactions: Graph neural networks and beyond", *Current Opinion in Systems Biology*, vol. 42, 100551, 2025.
 3. J. Lee, C.H. Nguyen, H. Mamitsuka, "Beyond rigid docking: deep learning approaches for fully flexible proteinligand interactions", *Briefings in Bioinformatics*, vol. 26 (5):bbaf454, 2025.
 4. M. Ishikawa, Z. Jiang, C.H. Nguyen, H. Hatsukawa, T. Hirai, H. Matsumoto, E. Saito, K. Okazaki, K. Endo, S. Terada, H. Mamitsuka, "Uncovering Key Features for Predicting Comorbid Chronic Eosinophilic Pneumonia in Chronic Rhinosinusitis via Machine Learning", *International Forum of Allergy & Rhinology*, e23613, 2025.
 5. M. Ishikawa, Z. Jiang, C.H. Nguyen, H. Hatsukawa, H. Mamitsuka, "Clinical Key Features Uncovered by Blood Eosinophilia-Based Machine Learning Classification of Chronic Rhinosinusitis", *International Forum of Allergy & Rhinology*, 2025.
 6. S.H. Nguyen, D.Q. Le, T.T. Nguyen, C.H. Nguyen, T.H. Ho, N.S. Vo, T. Nguyen, H.A. Nguyen, M.D. Cao, "PanKA: Leveraging population pangenome to predict antibiotic resistance" *iScience*, vol. 27 (9), 110623, 2024.
 7. D.Q. Le, T.A. Nguyen, S.H. Nguyen, T.T. Nguyen, C.H. Nguyen, H.T. Phung, T.H. Ho, N.S. Vo, T. Nguyen, H.A. Nguyen, M.D. Cao, "Efficient inference of large prokaryotic pangenomes with PanTA", *Genome Biology*, vol. 21 (1), 209, 2024.
 8. D.Q. Le, S.H. Nguyen, T.T. Nguyen, C.H. Nguyen, T.H. Ho, N.S. Vo, T. Nguyen, H.A. Nguyen, M.D. Cao, "AMRViz enables seamless genomics analysis and visualization of antimicrobial resistance", *BMC bioinformatics*, vol. 25(1), 193, 2024.

9. T. Cao, L. Sun, C.H. Nguyen, H. Mamitsuka, "Learning Low-Rank Tensor Cores with Probabilistic l0-Regularized Rank Selection for Model Compression", *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence (IJCAI-24)*, pp. 3780-3788, 2024.
10. D.Q. Le, T.T. Nguyen, C.H. Nguyen, T.H. Ho, N.S. Vo, T. Nguyen, H.A. Nguyen, M.D. Cao, S.H. Nguyen, "AMRomics: a scalable workflow to analyze large microbial genome collections", *BMC Genomics* vol. 25 (1), n.709, 2024.
11. V.H. Do, S.H. Nguyen, D.Q. Le, T.T. Nguyen, C.H. Nguyen, T.H. Ho, N.S. Vo, T. Nguyen, H.A. Nguyen, M.D. Cao, "Pasa: leveraging population pangenome graph to scaffold prokaryote genome assemblies", *Nucleic acids research*, vol. 52 (3), pp. e15, 2024.
12. X. Wang, L. Sun, C.H. Nguyen and H. Mamitsuka, "Multiplicative Sparse Tensor Factorization for Multi-View Multi-Task Learning", *Frontiers in Artificial Intelligence and Applications*, 372: *ECAI 2023* pp. 2560-2567, 2023.
13. D.A. Nguyen, C.H. Nguyen, H. Mamitsuka, "Central-Smoothing Hypergraph Neural Networks for Predicting Drug-Drug Interactions", *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35 (8), pp. 11620-11625, 2023.
14. D. A. Nguyen, C. H. Nguyen, H. Mamitsuka, "SPARSE: a sparse hypergraph neural network for learning multiple types of latent combinations to accurately predict drug-drug interactions", *Bioinformatics, (Proceedings of the 30th International Conference on Intelligent Systems for Molecular Biology (ISMB 2022))*, vol 38, pp i333-i341, 2022.
15. D. A. Nguyen, K. A. Nguyen, C. H. Nguyen, K. Than, "Boosting prior knowledge in streaming variational Bayes", *Neurocomputing*, vol. 424, pp.143-159, 2021.
16. C. H. Nguyen and H. Mamitsuka, "Learning on Hypergraphs with Sparsity", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, pp. 2710-2722, 2021.
17. M. Cai, C. H. Nguyen, H. Mamitsuka, L. Li, "XGSEA: CROSS-species Gene Set Enrichment Analysis via domain adaptation", *Briefings in Bioinformatics*, vo. 25, no. 5, 2021.
18. H. Kaneko et. al., "Eukaryotic virus composition can predict the efficiency of carbon export in the global ocean", *iScience*, vol. 24, no.1, 2021.
19. D. A., Nguyen, C. H., Nguyen and H. Mamitsuka, "A Survey on Adverse Drug Reaction Studies: Data, Tasks, and Machine Learning Methods", *Briefings in Bioinformatics*, vol. 22, no. 1, pp. 164-177, Jan 2021.
20. D. H. Nguyen, C. H. Nguyen, H. Mamitsuka, "Learning subtree pattern importance for Weisfeiler-Lehman based graph kernels", *Machine Learning* 110 (7), pp. 1585-1607, 2021.
21. C. H. Nguyen, "Structured Learning in Biological Domain", *Journal of Systems Science and Systems Engineering*, vol. 29, no. 4, pp 440-453, 2020.
22. L. Sun, C.H. Nguyen, H. Mamitsuka, "Fast and Robust Multi-View Multi-Task Learning via Group Sparsity", in *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI 2019)*, 3499-3505, 2019.
23. L. Sun, C.H. Nguyen, H. Mamitsuka, "Multiplicative Sparse Feature Decomposition for Efficient Multi-View Multi-Task Learning", in *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI 2019)*, 3506-3512, 2019.
24. D.H. Nguyen, C.H. Nguyen, H. Mamitsuka, "ADAPTIVE: leArning DAta-dePendenT, concise molecular VEctors for fast, accurate metabolite identification from tandem mass spectra". *Bioinformatics, (Proceedings of the 26th International Conference on Intelligent Systems for Molecular Biology (ISMB 2019))*, vol. 35, no. 14, pp. i164-i172, July 2019.
25. H., Seki, C.H., Nguyen, V.-N., Huynh, M., Inuiguchi (Eds.) *Integrated Uncertainty in Knowledge Modelling and Decision Making. Proceedings of 7th International Symposium, IUKM 2019*, Springer LNCS 11471, 2019.
26. N. Wicker, C.H. Nguyen, H. Mamitsuka, "A p-Laplacian random walk: application to video games". *Austrian Journal of Statistics*, 48(5), 11-16, 2019.
27. D.H. Nguyen, C.H. Nguyen, H. Mamitsuka, "Recent advances and prospects of computational methods for metabolite identification: a review with emphasis on machine learning approaches". *Briefings in Bioinformatics*, vol. 20, no. 6, pp. 2028-2043, Nov 2019.

28. D.H. Nguyen, C.H. Nguyen, H. Mamitsuka, "SIMPLE: Sparse Interaction Model over Peaks of MoLEcules for Fast, Interpretable Metabolite Identification from Tandem Mass Spectra". *Bioinformatics, (Proceedings of the 26th International Conference on Intelligent Systems for Molecular Biology (ISMB 2018))*, pp. i323-i332, 2018.
29. N. Wicker, C.H. Nguyen, H. Mamitsuka, "Some Properties of a Dissimilarity Measure for Labeled Graphs". *Publications Mathematiques de Besancon*, pp. 85-94, 2016.
30. A. Mohamed, C.H. Nguyen, H. Mamitsuka, "NMRPro: An integrated web component for interactive processing and visualization of NMR spectra." *Bioinformatics*, vol. 32, no. 13, pp. 2067-2068, 2016.
31. C.H. Nguyen, H. Mamitsuka, "New Resistance Distances with Global Information on Large Graphs". *The 19th International Conference on Artificial Intelligence and Statistics (AISTATS)*, May 2016.
32. A. Mohamed, C.H. Nguyen, H. Mamitsuka, "Current status and prospects of computational resources for natural product dereplication: A review" *Briefings in Bioinformatics*, vol. 17, no. 2, pp. 309-321, 2016.
33. A. Mohamed, T. Hancock, C.H. Nguyen, H. Mamitsuka, "NetPathMiner: R/Bioconductor package for network path mining through gene expression", *Bioinformatics*. vol. 30, no. 21, pp. 3139-3141, Nov 01, 2014.
34. C. H. Nguyen, N. Wicker and H. Mamitsuka, "Selecting Graph Cut Solutions via Global Graph Similarity", *IEEE Transactions on Neural Networks and Learning Systems*. vol. 25, no. 7, pp. 1407-1412, 2014.
35. N. Wicker, C. H. Nguyen and H. Mamitsuka, "A new dissimilarity measure for comparing labeled graphs", *Linear Algebra and its Applications*. vol. 438, no. 5, pp. 2331-2338. Mar 01, 2013.
36. C. H. Nguyen and H. Mamitsuka, "Latent Feature Kernels for Link Prediction on Sparse Graphs" *IEEE Transactions on Neural Networks and Learning Systems*. vol. 23, no. 11, pp. 1793-1804, 2012.
37. C. H. Nguyen and H. Mamitsuka, "Kernels for link prediction with latent feature models," *The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (2) (ECML/PKDD 2011)*, pp. 517-532, 2011.
38. C. H. Nguyen and H. Mamitsuka, "Discriminative graph embedding for label propagation," *IEEE Transactions on Neural Networks*, vol. 22, no. 9, pp. 1395-1405, 2011.
39. C. H. Nguyen, T. B. Ho, and V. Kreinovich, "Estimating quality of support vector machines learning under probabilistic and interval uncertainty: Algorithms and computational complexity," *Interval / Probabilistic Uncertainty and Non-Classical Logics*, pp. 57-69, 2008.
40. C. H. Nguyen and T. B. Ho, "An efficient kernel matrix evaluation measure," *Pattern Recognition*, vol. 41, no. 11, pp. 3366-3372, 2008.
41. H. Tanabe, T. B. Ho, C. H. Nguyen, and S. Kawasaki, "Simple but effective methods for combining kernels in computational biology," *RIVF*, pp. 71-78, 2008.
42. C. H. Nguyen and T. B. Ho, "Kernel matrix evaluation," *International Joint Conference on Artificial Intelligence (IJCAI2007)*, pp. 987-992, 2007.
43. T. B. Ho, C. H. Nguyen, S. Kawasaki, S. Q. Le, and K. Takabayashi, "Exploiting temporal relations in mining hepatitis data," *New Generation Computing*, vol. 25, no. 3, pp. 247-262, 2007.
44. T. B. Ho, C. H. Nguyen, S. Kawasaki, and K. Takabayashi, "Temporal relations extraction in mining hepatitis data," *The Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD2007)*, pp. 523-530, 2007.
45. C. H. Nguyen and T. B. Ho, "An imbalanced data rule learner," *The European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD2005)*, pp. 617-624, 2005.