

Hiroto Saigo

Associate Professor
Faculty of Information Science and Electrical
Engineering, Kyushu University

Visiting Researcher
RIKEN Center for Advanced Intelligence
Project (AIP)

Address: 744 Motoooka, Nishi-ku, Fukuoka 819-0395, Japan
Tel: +81 92 802 3783
Fax: +81 92 802 3783
Email: saigo@inf.kyushu-u.ac.jp
Homepage: <http://www.i.kyushu-u.ac.jp/%7Esaigo/>



Personal

Born on July 27, 1977.

Japanese Nationality.

Language: Japanese, English, German (intermediate).

Research Record

2020–present, Visiting Researcher, RIKEN Center for Advanced Intelligence Project (AIP)

2019, Visiting Associate Professor, University of Tokyo

2019, Visiting Associate Professor, Kyoto University

2016–present, Associate Professor, Faculty of Information Science and Electrical Engineering, Kyushu University

(Project) Manufacturing informatics. We are working on the prediction of the slag viscosity in the blast furnace, an important problem in steelmaking industry. Collaborative work with Research Center for Steel, Kyushu University.

(Project) Materials informatics. We are working on the problem of finding the optimal materials for electrode for the solid oxide fuel cells. Collaborative work with Institute for Carbon-Neutral Energy Research, Kyushu University.

2010–2015, Associate Professor, Department of Bioscience and Bioinformatics, Kyushu Institute of Technology, Japan.

(Project) Genome Wide Association Study. With special interest in identifying multiple factors that are working coordinately.

2008–2010, Research Scientist, Department of Computational Biology and Applied Algorithms, Max Planck Institute for Informatics, Saarbrücken, Germany.

(Project) Development and application of statistical learning methods to the problems associated with Human Immunodeficiency Virus (HIV). I am currently working on the prediction of the response of the latest regimen to a patient given all his/her past treatment history. Additionally, I have organized a machine learning seminar.

2006–2008, Research Scientist, Department of Empirical Inference, Max Planck Institute for Biological Cybernetics, Tübingen, Germany.

(Project) Development of machine learning methods and their applications to problems in bioinformatics and cheminformatics. I have proposed a graph-based virtual screening approach which integrates data mining and machine learning methods in (MLG 2006, Machine Learning 2009), and received the best paper award for this achievement.

2004–2006, Research Associate, Institute for Chemical Research, Kyoto University, Kyoto, Japan.

(Project) Development of kernel methods for detecting remote homology between protein sequences. The resulting kernel measures the similarity of proteins by considering all the possible alignment, and was successfully applied to classifying proteins into functional classes in combination with support vector machine. The resulting papers made a considerable impact on the field, and have been cited more than 200 times. (according to Google scholar)

2003–2004, Visiting Student, University of California Irvine, Irvine, USA (Advisor: Dr. Pierre Baldi).

(Project)

Construction of a server system for classifying proteins with disulphide bonds as a building block of protein structure prediction system.

Development of kernel methods for measuring the similarity between chemical compounds.

Discrimination of amino acid sequences of p53 tumor suppressor proteins using kernels for sequences.

Education

2003–2006, Doctor of Informatics, Informatics, Kyoto University, Kyoto, Japan.
Thesis: Local Alignment Kernels for Protein Homology Detection (Supervisor: Dr. Tatsuya Akutsu)

2003–2004, Visiting Student, University of California Irvine, Irvine, USA

2001–2003, Master of Informatics, Informatics, Kyoto University, Kyoto, Japan.

1997–2001, Bachelor of Engineering, Electrical and Electronics Engineering, Sophia University, Tokyo, Japan.

1993–1996, Sugamo High School, Tokyo, Japan.

Professional Activities

Program Committee of ACML2009, IBIS (2012,2019), PRIB (2012,2013), IIBMP (2013,2015,2020), GIW(2014,2018), BIBM (2011-2017), AAAI(2018-2020), NIPS(2016-2023), ICML(2017-2023), APBC2018, AISTATS(2016-2018)

Reviewer of NIPS (2010-2015), Bioinformatics, BMC Bioinformatics, Computational Intelligence in Bioinformatics, IEEE Transactions on Systems Man and Cybernetics (C), Machine Learning, Journal of Machine Learning Research (JMLR), IEICE Transactions on Information and Systems, Journal of Information Processing Society of Japan (IPSJ), Foundation for Medical Research in France

Grants

Kiban (B), *Head*, Research Grant from Japan Society for the Promotion of Science (2023-2028, Grant Number 23H03356).

Houga, *Head*, Research Grant from Japan Society for the Promotion of Science (2022-2025, Grant Number 22K19834).

Kiban (B), Research Grant from Japan Society for the Promotion of Science (2018-2020, Grant Number 21H01684).

Kiban (B), *Head*, Research Grant from Japan Society for the Promotion of Science (2018-2020, Grant Number 19H04176).

Kiban (B), Research Grant from Japan Society for the Promotion of Science (2018-2020, Grant Number 18H01762).

Wakate (A), *Head*, Research Grant from Japan Society for the Promotion of Science (2013-2017, Grant Number 25700004).

Wakate (B), *Head*, Research Grant from Japan Society for the Promotion of Science (2011-2013, Grant Number 23700338).

Houga, Research Grant from Japan Society for the Promotion of Science (2013-2015, Grant Number 25540062).

Wakate (B), Research Grant from Japan Society for the Promotion of Science (2011-2013, Grant Number 23710085).

Awards and Honors

Student's Best Poster Award in IIBMP 2014.

Incentive Award in Data Mining and Statistical Science 2007.

Best Paper Award in Mining and Learning with Graphs 2006.

Education Abroad Program Scholarship from Kyoto University 2003.

Interests

Materials Informatics, Manufacturing Informatics, Virtual screening, Genome-Wide Association Analysis (GWAS), Large scale hypothesis testing, Variable selection, Mining and Learning with structured data

Courses taught

Machine Learning (Graduate, 2011-present)
Bioinformatics (Undergrad, 2016-present)
Data Science (Undergrad, 2016-present)
Cyber Security (Undergrad, 2018-present)
Programming (Undergrad, 2010-2015,2017)
Data Structure and Algorithm (Undergrad, 2019-present)
Numerical Computation (Undergrad, 2010-2015)
Introduction to Bioinformatics (Undergrad, 2012-2015)
Numerical Computation Practice (Undergrad, 2014-2015))
Presentation and Discussion (Undergrad, 2010-2013)
Scientific English (Undergrad, 2015,2022)

Publications

Journal Articles (Citation counts are according to Google Scholar)

1. Pratyush, P., Pokharel, S., Saigo, H., KC.D.B.: *pLMSNOSite: an ensemble-based approach for predicting protein S-nitrosylation sites by integrating supervised word embedding and embedding from pre-trained protein language model*, *BMC Bioinform.* 24(1): 41 (02 2023) .
2. Saigo, H., Bahadur, K.C.D, Saito, N.:*Einstein-Roscoe regression for the slag viscosity prediction problem in steelmaking*, *Scientific Reports* 12, 6541 (04 2022), 2 citations.
3. Li, W., Saigo, H., Tong, B., Suzuki, E.:*Topic modeling for sequential documents based on hybrid inter-document topic dependency*. *J. Intell. Inf. Syst.* 56(3): 435-458 (2021), 4 citations.
4. Takayanagi, M., Tabei, Y., Suzuki, E., Saigo, H.: *Sparse nonnegative interaction models*, *IEEE Access*, Vol 9, pp. 109994-110005, (08 2021). DOI: 10.1109/ACCESS.2021.3099473.
5. Chaudhari, M., Thapa, N., S., Roy, K., Newman, R.H., Saigo, H., KC, D.B.: *DeepRMethylSite: a deep learning based approach for prediction of arginine methylation sites in proteins*, *Molecular Omics*, 16, 448, (10 2020), 11 citations.
6. Thapa, N., Chaudhari, M., McManus, S., Roy, K., Newman, R.H., Saigo, H., KC, D.B.: *DeepSuccinyl-Site: a deep learning based approach for protein succinylation site prediction*, *BMC Bioinformatics*, 21(Suppl 3):63, (04 2020), 34 citations.
7. Al-barakati, H.J., Thapa, N., Saigo, H., Roy, K., Newman, R.H., Bahadur, K.C.D.: *RF-MaloSite and DL-Malosite: Methods based on random forest and deep learning to identify malonylation sites*, *Computational and Structural Biotechnology Journal*, Vol 18. pp. 852-860, (02 2020), 0 citations.
8. Al-barakati, H.J., Saigo, H., Newman, R.H., Bahadur, K.C.D.: *RF-GlutarySite: a random forest predictor for glutarylation sites* , *Molecular Omics*, Vol 15. pp. 189-204, (04 2019), 26 citations.

9. White, C., Ismail, H., Saigo, H., Bahadur, K.C.D.: *CNN-BLPred: A Convolutional Neural Network based predictor for Beta-Lactamases (BL) and their classes*, *BMC Bioinformatics* 18(Suppl 16):577, 2017, 22 citations.
10. Kodama, K., Saigo, H.: *KDE: a Kernel-based approach to Detecting high-order genetic Epistasis*, *Journal of Bioinformatics and Computational Biology* 14(5), 1644003, 2016, 2 citations.
11. Suryanto, C. H., Saigo, H., Fukui, K.: *Protein Structure Comparison Based on 3D Molecular Visualization Images* *IEEE Transactions on Computational Biology and Bioinformatics*, Vol.PP, Issue:99 (08 2016), 11 citations.
12. Ismail, H.D., Saigo, H., Bahadur, K.C.D., *RF-NR: Random forest based approach for improved classification of Nuclear Receptors*, *IEEE Transactions on Computational Biology and Bioinformatics*, Vol.15, Issue:6 (11 2018), 7 citations.
13. Shao, Z., Hirayama, Y., Yamanishi, Y., Saigo, H.: *Mining discriminative patterns from graph data with multiple labels and its application to QSAR* *Journal of Chemical Information and Modeling*, 55(12), 2519-2527, (12 2015), 9 citations.
14. Saigo, H., Kashima, H., Tsuda, K.: *Fast iterative mining using sparsity-inducing loss functions*, *IEICE Transaction on Information and Systems*, Vol.E96-D No.8 pp.1766-1773 (08 2013), 0 citations.
15. Yamanishi, Y., Pauwels, E., Saigo, H., Stoven, V.: *Extracting sets of chemical substructures and protein domains governing drug-target interactions*, *Journal of Chemical Information and Modeling*, 51(5), 1183-1194, (05 2011). 85 citations.
16. Saigo, H., Altmann, A., Bogojeska, J., Mueller, F., Nowozin, S., and Lengauer, T.: *Learnig from past treatments and their outcome improves prediction of in vivo response to anti-HIV therapy*, *Statistical Applications in Genetics and Molecular Biology* 10(1) (01 2011). 12 citations.
17. Saigo, H., Hattori, M. Kashima, H. and Tsuda, K.: *Reaction graph kernels that predict EC numbers of unknown enzymatic reactions in the secondary metabolism of plant*, *BMC Bioinformatics* 11 (Supple 1), 1-7, (01 2010), 15 citations.
18. Saigo, H., Nowozin, S., Kadowaki, T., Kudo, T., and Tsuda, K.: *gBoost: A mathematical programming approach to graph classification and regression*, *Machine Learning* 75(1) 69-89, (04 2009), 149 citations.
19. Saigo, H., Uno, T. and Tsuda, K.: *Mining complex genotypic features for predicting HIV-1 drug resistance*, *Bioinformatics* 23(18), 2455-2462 (09 2007), 66 citations.
20. Saigo, H., Vert, J.-P. and Akutsu, T.: *Optimizing amino acid substitution matrices with a local alignment kernel*, *BMC Bioinformatics* 7(246), 1-12 (05 2006), 67 citations.
21. Danziger, S. A., Swamidass, S. J., Zeng, J., Dearth, L. R., Lu, Q., Cheng, J. H., Cheng, J. L., Hoang, V. P., Saigo, H., Luo, R., Baldi, P., Brachmann, R. K., and Lathrop, R. H.: *Functional census of mutation sequence spaces: The example of p53 cancer rescue mutants*, *IEEE Transactions on Computational Biology and Bioinformatics* 3(2), 114-125 (04 2006), 55 citations.
22. Cheng, J., Saigo, H. and Baldi, P.: *Large-scale prediction of disulphide bridges using kernel methods, two-dimensional recursive neural networks, and weighted graph matching*, *Proteins* 62(3), 617-629 (02 2006), 158 citations.
23. Matsuda, S., Vert, J.-P., Saigo, H., Ueda, N., Toh, H. and Akutsu, T.: *A novel representation of protein sequences for prediction of subcellular location using support vector machines*, *Protein Science* 14(11), 2804-2813 (2005), 181 citations.

24. Ralaivola, L., Swamidass, J. S., Saigo, H. and Baldi, P.: *Graph Kernels for Chemical Informatics, Neural Networks* 18(8), 1093-1110 (2005), 527 citations.
25. Saigo, H., Vert, J.-P., Ueda, N. and Akutsu, T.: *Protein homology detection using string alignment kernels, Bioinformatics* 20(11), 1682-1689 (2004), 482 citations.

Proceedings (Citation counts are according to Google Scholar)

1. Yafune, R., Sakuma, D., Tabei, Y., Saito, N., Saigo, H.: *Automatically mining relevant variable interactions via sparse Bayesian learning; International Conference of Pattern Recognition (ICPR2020)*, Milan, January 15-20, 2021.
2. Li, W., Matsukawa, T., Saigo, H., Suzuki, E.: *Context-Aware Latent Dirichlet Allocation for Topic Segmentation; Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2020)*, Singapore, May 14, 2020, 4 citations.
3. Thapa, N., Chaudhari, M., McManus, S., Roy, K., Newman, R.H., Saigo, H., KC, D.B.: *DeepSuccinyl-Site: a deep learning based approach for protein succinylation site prediction, MCBIOS*, Birmingham, March 28-30, 2019, 34 citations.
4. Albarakati, H., Saigo, H., Newman, R.H., KC, D.B.: *SVM-GlutarySite: A support vector machine-based prediction of Glutarylation sites from protein sequences, Joint GIW/ABACBS-2019 Bioinformatics Conference*, Sydney, December 9-11, 2019, 0 citations.
5. Takayanagi, M., Tabei, Y., Saigo, H.: *Entire regularization path for sparse nonnegative interaction model, International Conference of Data Mining (ICDM2018)*, Singapore, November 17-20, 2018, 3 citations.
6. White, C., Ismail, H., Saigo, H., Bahadur, K.C.D.: *CNN-BLPred: A Convolutional Neural Network based predictor for Beta-Lactamases (BL) and their classes, International Conference on Bioinformatics (InCoB2017)*, Shenzhen, China, September 20-22, 2017, 22 citations.
7. Kodama, K., Saigo, H.: *KDSNP: a Kernel-based approach to Detecting high-order genetic Epistasis, International Conference on Genome Informatics (GIW2016)*, Shanghai, China, October 3-5, 2016, 2 citations.
8. Tabei, Y., Saigo, H., Yamanishi, Y., Puglisi, S., *Scalable Partial Least Squares Regression on Grammar-Compressed Data Matrices, In Proceedings of the 22nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD2016)*, (8 2016), 28 citations.
9. Ismail, H.D., Saigo, H., Bahadur, K.C.D., *RF-NR: Random forest based approach for improved classification of Nuclear Receptors, In Proceedings of the International Conference on Genome Informatics & International Conference on Bioinformatics (GIW/InCoB2015)*, (9 2015), 7 citations.
10. Suryanto, C.H., Saigo, H., Fukui, K.: *Protein Clustering on Grassmanns Manifold, Pattern Recognition in Bioinformatics (PRIB2012)*, Tokyo, Japan. (11 2012), 7 citations.
11. Saigo, H., Hattori, M. Kashima, H. and Tsuda, K.: *Reaction graph kernels that predict EC numbers of unknown enzymatic reactions in the secondary metabolism of plant, In Proceedings of the eighth Asia Pacific Bioinformatics Conference (APBC2010)*, Bangalore, India, (1 2010). 15 citations
12. Chiappa, S., Saigo, H. and Tsuda, K.: *A Bayesian Approach to Graph Regression with Relevant Subgraph Selection In Proceedings of the Siam International Conference on Data Mining (SDM2009)*, Nevada, U.S. (4 2009), 6 citations.
13. Saigo, H. and Tsuda, K.: *Iterative Subgraph Mining for Principal Component Analysis In Proceedings of the 8th IEEE International Conference on Data Mining (ICDM2008)*, 1007-1012, Pisa , Italy. (12 2008). 12 citations.

14. Kashima, H., Yamasaki, K., Saigo, H. and Inokuchi, A.: *Regression with Interval Output Values*, In *Proc. 19th International Conference on Pattern Recognition (ICPR2008)*, (2008), 5 citations.
15. Saigo, H., Krämer, N. and Tsuda, K.: *Partial Least Squares Regression for Graph Mining*, In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD2008)*, 578-586, Las Vegas, U.S. (08 2008), 106 citations.
16. Saigo, H., Hattori, M. and Tsuda, K.: *Reaction graph kernels for discovering missing enzymes in the plant secondary metabolism*, *NIPS Workshop on Machine Learning in Computational Biology*, Whistler, BC, Canada (12 2007), 0 citations.
17. Kashima, H., Yamazaki, K., Saigo, H. and Inokuchi, A.: *Regression with Intervals*, *International Workshop on Data-Mining and Statistical Science (DMSS2007)*, Tokyo, Japan (10 2007), 0 citations.
18. Saigo, H., Kadowaki, T., Kudo, T. and Tsuda, K.: *Graph boosting for molecular QSAR analysis*, *NIPS Workshop on Machine Learning in Computational Biology*, Whistler, BC, Canada (12 2006), 0 citations.
19. Saigo, H., Kadowaki, T. and Tsuda, K.: *A Linear Programming Approach for Molecular QSAR analysis*, *International Workshop on Mining and Learning with Graphs (MLG2006)*, 85-96, Berlin, Germany (9 2006) **Best Paper Award**. 26 citations

Book Chapters (Citation counts are according to Google Scholar)

1. Saigo, H. and Tsuda, K.: *Matrix Decomposition-based Dimensionality Reduction on Graph Data*, In Sakr, S. and Pardede, E. editors *Graph Data Management: Techniques and Applications* IGI Global, (2011), 0 citations.
2. Saigo, H., and Tsuda, K. *Graph Mining in Chemoinformatics*, In Huma Lodhi and Yoshihiro Yamanishi editors, *Chemoinformatics and Advanced Machine Learning Perspectives: Complex Computational Methods and Collaborative Techniques*, IGI global, (2010), 2 citations.
3. Kashima, H., Saigo, H., Hattori, M., and Tsuda, K. *Graph Kernels for Chemoinformatics*, In Huma Lodhi and Yoshihiro Yamanishi editors, *Chemoinformatics and Advanced Machine Learning Perspectives: Complex Computational Methods and Collaborative Techniques*, IGI global, (2010), 18 citations.
4. Tsuda, K. and Saigo, H. *Graph Classification*, 337-364, In Charu C. Aggarwal and Haixun Wang editors, *Managing and Mining Graph Data*, Springer, New York, NY, (2010), 26 citations.
5. Vert, J.-P., Saigo, H., and Akutsu, T.: *Local Alignment Kernels for Biological Sequences*, 131-153, In Bernhard Schölkopf, Koji Tsuda and Jean-Philippe Vert editors, *Kernel Methods in Computational Biology*, MIT Press, Cambridge, MA, (2004), 91 citations.

Short-time research stay and Talks

1. Bad Neuenahr, Germany (Japan America Germany Frontiers of Science Symposium, 2017.09)
2. National Yang-Ming University, Taipei, Taiwan (Talk title: Mining and learning with structured data at BIT2016), 2016.03.
3. Kurume University, Fukuoka, Japan (Host: Koji Yabara, Talk title: Mining Discriminative Patterns from Graph Data with Multiple Labels and its Application to Quantitative Structure-Relationship Analysis (QSAR)), Jan, 2016.
4. Kagoshima University, Kagoshima, Japan (Host: Tatsuhiko Furukawa, Talk title: Introduction to Chemoinformatics), Jan, 2015.

5. Tohoku University, Sendai, Japan (Host: Masayoshi Hatayama, Talk title: Pteris Vittata analysis report), Feb, 2013.
6. Friedrich Mischer Laboratory, Tuebingen, Germany (Host: Karsten Borgwardt, Talk title: Multiple response regression for graph mining), Nov, 2013.
7. Imperial College London, U.K. (Host: Marc Deisenroth, Talk title: Multiple response regression for graph mining), Oct, 2013.
8. Novartis, Basel, Switzerland (Host: Jean-Luc Perret, Talk title: Clustering approach to drug discovery, Dec, 2012.
9. Ecole des Mines de Paris and Paris Tech in Paris, France (Host: Yoshihiro Yamanishi, Talk title: Learning from past treatments and their outcome improved prediction of in vivo response to anti-HIV therapy), Feb 2010.
10. Eidgenoesische Technische Hochschule (ETH) Zuerich, Basel, Switzerland (Host: Niko Beerenwinkel, Talk title: Incorporating detailed information on treatment history improves prediction of response to anti-HIV therapy), Dec 2009.
11. Universitet van Amsterdam, Amsterdam, Netherland (Host: Goehan Ertaylan, Talk title: Incorporating detailed information on treatment history improves prediction of response to anti-HIV therapy), Dec 2009.
12. Ecole des Mines de Paris and Paris Tech in Paris, France (Host: Jean-Philippe Vert, Talk title: Partial Least Squares Regression for Graph Mining), May 2008.
13. Fraunhofer FIRST, Berlin, Germany (Host: Motoaki Kawanabe, Talk title: A Linear Programming Approach for Molecular QSAR analysis), September 2006.
14. Ecole des Mines de Paris, Paris, France (Host: Jean-Philippe Vert), May 2005.
15. Information-Based Induction Science (IBIS) 2005, Waseda University, Tokyo, Japan (Host: Koji Tsuda, Talk title: Classification of chemical compounds using graph kernels), Oct 2005.
16. Ecole des Mines de Paris, Paris, France (Host: Jean-Philippe Vert), April 2003.