Bioinformatics Center
– Bio-knowledge Engineering –

https://www.bic.kyoto-u.ac.jp/pathway/index.html

Prof
MAMITSUKA, Hiroshi
(D Sc)

Senior Lect
NGUYEN, Hao Canh
(D Knowledge Science)

Guest Res Assoc*
PETSCHNER, Peter
(Ph D)

*Semmelweis University, Hungary, 28 November 2020–15 November 2022

Students

NGUYEN, Duc Anh (D3)

Guest Scholars

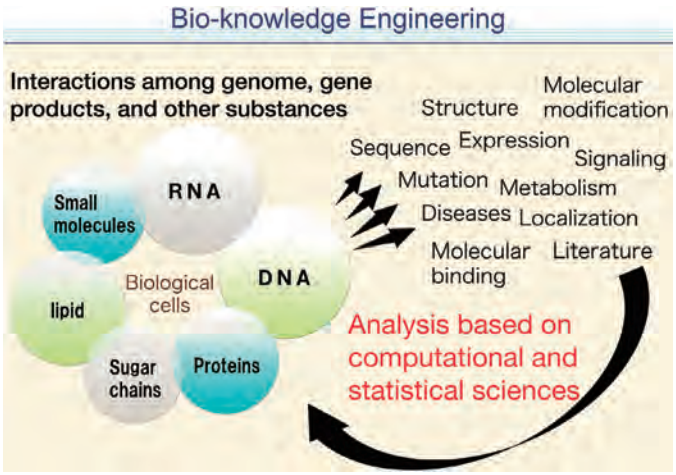
KLIPP, Edda (Ph D)	Humboldt University Berlin, Germany, 12 April 2022–15 May 2022
RAUN, Rune Linding (Ph D)	Humboldt University Berlin, Germany, 22 April 2022–6 May 2022
MALHI, Avleen Kaur (Ph D)	Bournemouth University, UK, 2 June 2022–18 August 2022

Guest Res Assoc

LI, Yufei	Xi'an Jiaotong University, China, 2 May 2022–23 April 2023
GOLDENBOGEN, Björn	Humboldt University Berlin, Germany, 4 April 2022–20 May 2022
GECSE, Kinga	Semmelweis University, Hungary, 13 August 2022–22 October 2022

Scope of Research

We are interested in graphs and networks in biology, chemistry, and medical sciences, including metabolic networks, protein-protein interactions and chemical compounds. We have developed original techniques in machine learning and data mining for analyzing these graphs and networks, occasionally combining with table-format datasets, such as gene expression and chemical properties. We have applied the techniques developed to real data to demonstrate the performance of the methods and find new scientific insights.



KEYWORDS

Bioinformatics
Machine Learning
Data Mining
Artificial Intelligence
Systems Biology

Recent Selected Publications

Nguyen, D. A.; Nguyen, C. H.; Petschner, P.; Mamitsuka, H., SPARSE: A Sparse Hypergraph Neural Network for Learning Multiple Types of Latent Combinations to Accurately Predict Drug-Drug Interactions, *Bioinformatics*, **38** *(Supplement 1) (Proceedings of the 30th International Conference on Intelligent Systems for Molecular Biology (ISMB 2022))*, i333-i341 (2022).

You, R.; Qu, W.; Mamitsuka, H.; Zhu, S., DeepMHCII: A Novel Binding Core-Aware Deep Interaction Model for Accurate MHC II-peptide Binding Affinity Prediction, *Bioinformatics*, **38** *(Supplement 1) (Proceedings of the 30th International Conference on Intelligent Systems for Molecular Biology (ISMB 2022))*, i220-i228 (2022).

Nguyen, C.H.; Mamitsuka, H., Learning on Hypergraphs with Sparsity, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **43,** 2710-2722 (2021).

Nakamura, A.; Takigawa, I.; Mamitsuka, H., Efficiently Enumerating Substrings with Statistically Significant Frequencies of Locally Optimal Occurrences in Gigantic String, *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI 2020)*, **34(4),** 5240-5247 (2020).

Nguyen, D. H.; Nguyen, C. H.; Mamitsuka, H., ADAPTIVE: learning DAta-dePendenT, concIse molecular VEctors for fast, accurate metabolite identification from tandem mass spectra, *Bioinformatics*, **35(14),** *(Proceedings of the 27th International Conference on Intelligent Systems for Molecular Biology (ISMB/ECCB 2019))*, i164-i172 (2019).

An Advanced Deep Learning Model for Predicting Side Effects of Drug-Drug Interactions

A drug-drug interaction (DDI) is a reaction between two drugs, whereby the effects of one drug are modified by the concomitant use of the second drug. A DDI might cause side effects, which are unwanted effects and are responsible for significant patient morbidity and mortality. Hence predicting side effects of drug-drug interactions is a very important task in pharmacology. Given known DDI data, a prominent approach is to use machine learning models to obtain highly accurate and fast DDI predictions.

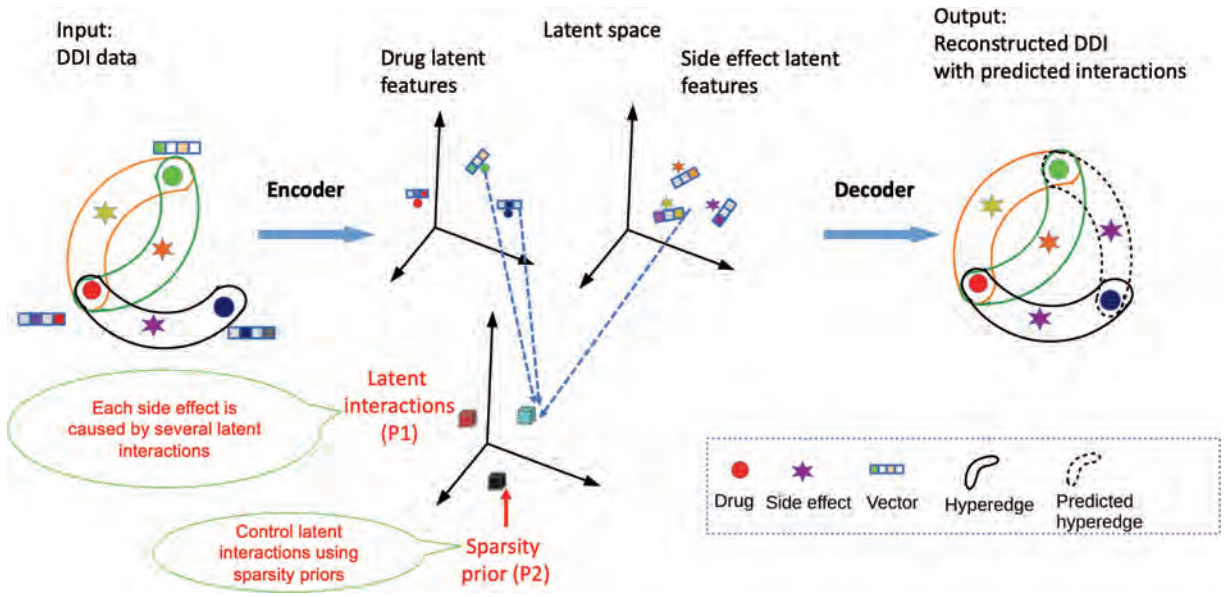
Recent state-of-the-art models assume that each side effect is caused by a unique combination of properties of the corresponding interacting drugs. However, there still exist two remaining problems that have not been addressed: (P1) in reality, a side effect might have multiple mechanisms that cannot be represented by a single combination of drug properties, and (P2) DDI data is sparse with a very few percentages of known DDIs, which might impair the model performances.

To address the above two problems, we proposed SPARSE, an advanced deep learning model for predicting drug-drug interactions with an illustration as in Figure 1. For problem (P1), we assumed that each side effect is caused by different combinations of drug properties. For problem (P2), we used a sparsity control to guide the model to fit with the sparse data.

We conducted experiments on the three real-world DDI benchmark datasets to show the advantage of the prediction performance of our proposed method. We compared SPARSE with other state-of-the-art methods under two commonly used measures: AUC (area under the ROC curve) and AUPR (area under the precision-recall curve). The results were shown in Figure 2. The figure shows that our method (SPARSE) outperformed other methods in both AUC and AUPR, suggesting that our method was suitable for drug-drug interaction prediction.

Reference

Duc Anh Nguyen, Canh Hao Nguyen, Peter Petschner, Hiroshi Mamitsuka, SPARSE: a sparse hypergraph neural network for learning multiple types of latent combinations to accurately predict drug–drug interactions, Bioinformatics, Volume 38, Pages i333-i341, https://doi.org/10.1093/bioinformatics/btac250.



Method	TWOSIDES		CADDDI		JADERDDI	
	AUC	AUPR	AUC	AUPR	AUC	AUPR
MRGNN	0.8452 ± 0.0036	0.8029 ± 0.0039	0.9226 ± 0.0015	0.7113 ± 0.0031	0.9049 ± 0.0009	0.3698 ± 0.0019
Decagon	0.8639 ± 0.0029	0.8094 ± 0.0024	0.9132 ± 0.0014	0.6338 ± 0.0029	0.9099 ± 0.0012	0.4710 ± 0.0027
SpecConv	0.8785 ± 0.0025	0.8256 ± 0.0022	0.8971 ± 0.0055	0.6640 ± 0.0014	0.8862 ± 0.0025	0.5162 ± 0.0047
HPNN	0.9044 ± 0.0003	0.8410 ± 0.0007	0.9495 ± 0.0004	0.7020 ± 0.0018	0.9127 ± 0.0004	0.5198 ± 0.0016
SBM	0.9337 ± 0.0002	0.8583 ± 0.0004	0.9588 ± 0.0006	0.8170 ± 0.0008	0.9428 ± 0.0006	0.5963 ± 0.0018
CentSmoothie	0.9348 ± 0.0002	0.8749 ± 0.0013	0.9846 ± 0.0001	0.8230 ± 0.0019	0.9684 ± 0.0004	0.6044 ± 0.0025
SPARSE	**0.9524** ± 0.0001	**0.8820** ± 0.0002	0.9837 ± 0.0010	**0.8843** ± 0.0012	**0.9698** ± 0.0008	**0.7348** ± 0.0018

Figure 2. Comparisons on prediction performance.