Mijin Kwon

Ph.D., Data Scientist

Samsung Electronics (Device Solution)

Korea Advanced Institute of Science and Technology (KAIST)

E-mail: mijinkwon924@gmail.com

Phone: +82 10-2871-3168

Google Scholar: [Click here](#)

LinkedIn: [Click here](#)

WORK EXPERIENCE

Data scientist	Samsung Electronics (Device Solution)	Oct. 2019 – Present
Systems architect	Samsung Electronics (Device Solution)	Oct. 2019 – Present
Data architect	Samsung Electronics (Device Solution)	Oct. 2019 – Present

EDUCATION

Ph.D.	Bio and Brain Engineering KAIST	3.94 / 4.3	Sept. 2015 – Aug. 2019
MSc.	Bio and Brain Engineering KAIST	3.95 / 4.3	Sept. 2013 – Aug. 2015
BSc.	Electronics Engineering Kyungpook National University	4.1 / 4.3	Mar. 2009 – Aug. 2013

RESEARCH AREAS

Big data mining, Data analysis, Data platform, Data architecture, Data visualization, Client-Server architecture, Database systems, Data lake, Data warehouse

RESEARCH & WORK EXPERIENCES & INTERESTS

[2019 ~ 2022]

Data Science

- **Big data analysis:** Machine learning or statistics based large-scale data analysis such as 1) regression model based prediction of wafer yields, 2) statistical model-based detection of commonality for low-yield wafers, 3) ARIMA/RNN/LSTM based predicting time-series trends of equipment sensors, 4) Wavelet-based large-scale time-series data transform/compression for improving efficiency in the data store and retention, 5) Deep learning-based automatic converting from natural language to SQL

  ▶ Required skills: Machine learning, Statistics, Python (scikit-learn, torch, keras, tensorflow), Visualization, DBMS, Hadoop, Impala, Spark

Systems Design

- **Client-Server Systems Design:** Defining the architecture of computer systems that help large-scale visualization or analysis of equipment data in an efficient and stable way so as to meet the business needs (# of users: 4,000)

  ▶ Required skills: Client-Server Model, Hardware and Software components, UI (C#), DBMS (PostgreSQL, NoSQL, Columnar DB), Multi-thread processing, Large-scale visualization, Network

Data architecture

- **Big data architecture:** Designing data architecture principles and data life cycle of equipment data produced in real-time (1TB per day) with the understanding of business requirements

  ▶ Required skills: Data lake, Database systems, Data warehouse, Data mart, Hadoop ecosystems, ETL, Kafka, Nifi, Data API

[2013 ~ 2019]

Systems biology

- **Next-generation sequencing:** Whole genome and RNA sequencing to capture individual-specific mutations or perturbations
- **Multi-omics data:** Integrative analysis of multi-omics datasets including genome, epigenome, transcriptome, proteome, microbiome for more exact finding significant markers of individuals
- **Large-scale biological network:** Individual-specific biological network creation and analysis
- **Personalized medicine:** disease prediction and drug sensitivity prediction based on biological networks and machine learning algorithms such as deep learning

LIST OF PUBLICATIONS

First-author publications
1. **Mijin Kwon**, Jinmyung Jung, Hasun Yu, and Doheon Lee*. "HIDEEP: a systems approach to predict hormone impacts on drug efficacy based on effect paths." Scientific reports 7, no. 1 (2017): 16600. (SCI) (30 Nov. 2017) (ISSN 2045-2322)
2. **Mijin Kwon,** Soorin Yim, Gwangmin Kim, Saehwan Lee, Chungsun Jeong, Doheon Lee*, "CODA-ML: context-specific biological knowledge representation for systemic physiology analysis." BMC bioinformatics 20.10 (2019): 248.
3. **Mijin Kwon,** Soorin Yim, Gwangmin Kim, Doheon Lee*, "Integrated Network-Based Computational Analysis for Drug Development", 2021, Springer, Recent Advances in Biological Network Analysis (Chapter 8) (ISBN 978-3-030-57172-6)
4. **Mijin Kwon,** Youngmin Woo, Woochang Hwang, Doheon Lee*, **"**Patient-specific sensitizer discovery to overcome drug resistance", 2019, *to be submitted soon*

Co-author publications
1. Jinmyung Jung, Yeeok Kang, Hyojung Paik, **Mijin Kwon**, Hasun Yu, Doheon Lee*, and Janet Kelso*. "Deconvoluting essential gene signatures for cancer growth from genomic expression in compound-treated cells." Bioinformatics (2018). (SCI) (1 Sept. 2018) (ISSN 1460-2059)
2. Jinmyung Jung, **Mijin Kwon**, Sunghwa Bae, Soorin Yim, and Doheon Lee*. "Petri net-based prediction of therapeutic targets that recover abnormally phosphorylated proteins in muscle atrophy." BMC systems biology 12, no. 1 (2018): 26. (SCI) (5 Mar. 2018) (ISSN 1752-0509)
3. Hasun Yu, Jinmyung Jung, Seyeol Yoon, **Mijin Kwon**, Sunghwa Bae, Soorin Yim, Jaehyun Lee, Seunghyun Kim, Yeeok Kang, and Dohoen Lee*. "CODA: integrating multi-level context-oriented directed associations for analysis of drug effects." Scientific reports 7, no. 1 (2017): 7519. (SCI) (08 Aug. 2017) (ISSN 2045-2322)
4. Woochang Hwang, Jaejoon Choi, **Mijin Kwon**, and Doheon Lee*. "Context-specific functional module based drug efficacy prediction." BMC bioinformatics 17, no. 6 (2016): 275. (SCI) (28 July 2016) (ISSN: 1471-2105)
5. Seyeol Yoon, Jinmyung Jung, Hasun Yu, **Mijin Kwon**, Sungji Choo, Kyunghyun Park, Dongjin Jang, Sangwoo Kim, and Doheon Lee*. "Context-based resolution of semantic conflicts in biological pathways." BMC medical informatics and decision making 15, no. 1 (2015): S3. (SCI) (20 May 2015) (ISSN: 1472-6947)

Conference papers
1. Woochang Hwang, Jaejoon Chio, **Mijin Kwon**, Doheon Lee*. "Cell line specific method based on functional modules for drug sensitivity prediction." In Proceedings of the ACM Ninth International Workshop on Data and Text Mining in Biomedical Informatics, pp. 29-29. ACM, 2015.

2. Hasun Yu, Jinmyung Jung, Seyeol Yoon, **Mijin Kwon**, Yeeok Kang, Sunghwa Bae, Doheon Lee*. "Development of a Framework for Constructing a Virtual Physiological Human with the Integration of COntext Specific Directed Associations (CODA)." In Proceedings of the ACM Ninth International Workshop on Data and Text Mining in Biomedical Informatics, pp. 18-18. ACM, 2015.
3. Jinmyung Jung, Hasun Yu, Seyeol Yoon, **Mijin Kwon**, Sungji Choo, Sangwoo Kim, Doheon Lee*. "Construction of Multi-level Networks Incorporating Molecule, Cell, Organ and Phenotype Properties for Drug-induced Phenotype Prediction." In Proceedings of the ACM 8th International Workshop on Data and Text Mining in Bioinformatics, pp. 47-47. ACM, 2014.
4. Seyeol Yoon, Jinmyung Jung, Hasun Yu, **Mijin Kwon**, Sungji Choo, Kyunghyun Park, Dongjin Jang, Sangwoo Kim, Doheon Lee*. "Systematic identification of context-dependent conflicting information in biological pathways." In Proceedings of the ACM 8th International Workshop on Data and Text Mining in Bioinformatics, pp. 9-9. ACM, 2014.

PATENTS

DOMESTIC
1. **Mijin Kwon**, "Method and system for predicting interaction between hormone and drug" (Application Number: 1020180034717).
2. Sunyoung Yoo, Minji Jung, **Mijin Kwon**, "Apparatus for Electronic Medical Record Providing" (Application Number: 1020150003455)

HONORS & AWARDS

1.	The best prize in KAIST Research Convergence Contest.	2014
2	The best poster prize in ISMB 2019 Conference	2019
3	Award for contribution on improving wafer yields based on M/L model, Samsung Electronics	2020
4	Award for contribution on urgent system support for equipment breakdown due to Austin blackout	2021
5	Award for contribution on decreasing equipment fault by creating C/S computer systems, Samsung Electronics	2021

GRANTS

Full tuition scholarships

The national scholarship fund for bachelor's programs	2009 – 2013
The national scholarship fund for master's programs	2013 – 2015
The national scholarship fund for doctoral programs	2015 – 2019

ACADEMIC PRESENTATIONS

1. **Mijin Kwon,** Soorin Yim, Gwangmin Kim, Saehwan Lee, Chungsun Jeong, Doheon Lee*, "CODA-ML: Context-Specific Biological Knowledge Representation for Systemic Physiology Analysis", ACM 12th International Workshop on Data and Text Mining in Biomedical Informatics (DTMBio), 22-26 Oct. 2018, Turin, Italy. (oral presentation)
2. **Mijin Kwon**, Doheon Lee, "Discovering Sensitizers to Overcome Cancer Drug Resistance", the 22nd Annual Conference on Research in Computational Molecular Biology (RECOMB), 21-24 Apr. 2018, Paris, France. (poster presentation)
3. **Mijin Kwon**, Doheon Lee, "Predicting efficacy and side effects of pharmaceutical compounds with a virtual human system CODA", The Milner Therapeutics Institute Annual Symposium, 2 Oct. 2017, Cambridge, United Kingdom. (oral presentation)
4. **Mijin Kwon** , Yousang Jo , Sunghwa Bae , Soorin Yim , Gwangmin Kim, Doheon Lee, "Analyzing biological processes of anatomical context-specific molecular networks through large-scale integration", The 25th Annual Conference Intelligent Systems for Molecular Biology (ISMB), 21-25 July 2017, Prague, Czech Republic. (poster presentation)
5. **Mijin Kwon**, Jinmyung Jung, Hasun Yu, Doheon Lee, "Network-based inference of the effects of hormones on drug efficacy", the Ninth Annual RECOMB/ISCB Conference on Regulatory & Systems Genomics, with DREAM Challenges & Cytoscape Workshop, 6-9 Nov. 2016, Pheonix, Arizona. (poster presentation)

TEACHING & MENTORING EXPERIENCE

Teaching assistance,   2013 – Present
Department of Bio and Brain Engineering, KAIST
Computer science or systems biology related undergraduate/graduate courses, covering the following topics: programming, machine learning algorithms, statistics, database, biological network analysis, data mining

Mentoring,   2015 –Present
Department of Bio and Brain Engineering, KAIST
Mentoring and counseling for undergraduate and graduate students

SKILLS

Technical skills
Programming languages: Python, R, SQL, C/C++

Language ability
English: Proficient (TOEIC: 925)

PROFESSIONAL SERVICE

Peer-reviewed articles for:
1. The Fifth International Workshop on Computational Network Biology: Modeling, Analysis, and Control (CNB-MAC 2018)
2. The 17th Asia Pacific Bioinformatics Conference (APBC 2019)

REFERENCES

Prof. Doheon Lee   **Prof. Kwanghyung Lee**
Department of Bio and Brain Engineering, KAIST
Email: dhlee@kaist.ac.kr   Email: khlee@kaist.ac.kr
Phone: +82-42-350-4316   Phone: +82-42-350-4313