CURRICULUM VITAE

October 2024

NAME: Yejin Kim

PRESENT TITLE:
Assistant Professor (full-time tenure track)

WORK ADDRESS:
Suite UCT E.760E
7000 Fannin Str.
Houston, TX 77030

UNDERGRADUATE EDUCATION:
2008 – 2012 BS in Industrial Engineering (Magna cum laude)
Pohang University of Science and Technology, Pohang, South Korea

GRADUATE EDUCATION:
2012 – 2017 Ph.D. in Computer Science
Pohang University of Science and Technology, Pohang, South Korea
Thesis: Interpretable and Federated Tensor Factorization for Computational Phenotyping
Advisor: Dr. Hwanjo Yu

POSTGRADUATE TRAINING
2017 – 2018 Postdoctoral Scientist
Pohang University of Science and Technology, Pohang, South Korea
Research on *machine learning for healthcare*
Mentor: Dr. Hwanjo Yu

ACADEMIC & ADMINISTRATIVE APPOINTMENTS:
2018 – Assistant Professor, McWilliams School of Biomedical Informatics, University of Texas Health Science Center at Houston (UTHealth), Texas
2018 – Founding Member, Center for Secure Artificial Intelligence for Healthcare, McWilliams School of Biomedical Informatics (MSBMI), UTHealth, Texas

PROFESSIONAL ORGANIZATIONS (AND COMMITTEES OF THESE):
2010 – 2017 Member, National Academy of Engineering of Korea Young Engineer Honor Society
2018 – Member, American Medical Informatics Association (AMIA)
2019 – Member, Association for Computing Machinery SIGKDD
2021 – Member, Institute for Stroke and Cerebrovascular Disease, UTHealth, Texas
2022 – Member, Ajou Global Fellow

HONORS AND AWARDS:
2008 – 2012 University full scholarship, Pohang University of Science and Technology
2020 Rank 2nd place (out of 54) in the International DREAM Challenge for Drug Repurposing Competition
2020 Selected as Highlight Poster in Alzheimer's Disease Association International Conference 2020 Technology and Dementia
2020 Microsoft Research Faculty Gift
2021 Selected as the top 10% most cited PLoS One paper in 2016
2022 Public media coverage by CNN, NPR, and ScienceDaily; invited to testify before U.S. Department of Health & Human Service for Alzheimer's claim data study

2023 Public media coverage by Yahoo Finance

SERVICE ON GRANT REVIEW PANELS, STUDY SECTIONS, COMMITTEES:

2020 Luxembourg National Science Foundation, Industrial Fellowship, Luxembourg
2022 Fonds de recherche du Québec, Ministry of Innovation, Science & Technology – State of Israel, Bilateral Research Cooperation Program MOST-FRQS, Canada, Israel
2023 UK Research and Innovation, Clinician Scientist Fellowship Referee, United Kingdom
2023 NIH CSR CDMA (Clinical Data Management and Analysis) Special Emphasis Panel ZRG1 HSS-P
2024 NIH CSR HSS (Health Services & System) ZRG1 HSS-B (02) M
2024 NIH NINDS Team Science COMBINE RM1 Special Panel Review ZNS1 SRB-S (21)
2024 NIH CSR MABS (Modeling and Analysis of Biological System) Study Section

SERVICE ON THE UNIVERSITY OF TEXAS HEALTH SCIENCE CENTER AT HOUSTON and SCHOOL OF BIOMEDICAL INFORMATICS COMMITTEES:

2019 Faculty Development Committee
2019 Students and Alumni Committee
2021 Fall 2021 Poster Session Judge
2022 Admission, Progression, and Graduation Committee
2022 Spring 2022 Poster Session Judge
2022 PhD Dissertation Committee member
Ivan Coronado, Kohong Lin, Tanjida Kabir, Xiaotian Ma, Dulin Wang, Yaobin Ling, Mehmet Enes Inam, Kaichen Tang, Rastko Stojsin
2023 PhD Dissertation Committee
Alan Pan, Qidi Xu, Yao-Shun Chung
2023 Admission, Progression, and Graduation Committee
2023 Departmental Chair Search Committee
2023 Interfaculty Council member (UTHealth)
2023 Representatives for Burks Scholarship (UTHealth)
2023 Chair, Faculty Search Committee, in the Center for Secure Artificial Intelligence for Healthcare.
2024 PhD Qualifying Exam Committee
2024 PhD Dissertation Committee
Christina Van Hal, Kunhee Kim

SERVICE TO THE LOCAL COMMUNITY:

2019 Co-organizer, 1st MSBMI Machine Learning Datathon, UTHealth
2020 Co-organizer, 2nd MSBMI Machine Learning Datathon, UTHealth
2020 Co-organizer, COVID-19 Houston Datathon, UTHealth
2021 Co-organizer, 3rd MSBMI Machine Learning Datathon, UTHealth
2021 Poster judge, 2021 Aging and Health Informatics Conference, School of Nursing & School of Information, University of Texas at Austin

SERVICE TO NATIONAL/INTERNATIONAL COMMUNITY

2017 Program committee for ACM Conference on Bioinformatics, Computational Biology and Health Informatics
2017-2018 Program committee for AMIA Annual Symposium
Program committee for World Wide Web Conference (WWW)
2019 Committee member, Cerner/SBMI National Data Science Challenge

2019- Program committee for AAAI Conference on Artificial Intelligence (AAAI)
2019- Program committee for IEEE BIBM International Conference on Bioinformatics and Biomedicine
2019- Program committee for SIGKDD Applied Data Science for Healthcare
2020 Program committee for WSDM Web Search and Data Mining Health Day
2020- Program committee for AMIA Knowledge Discovery and Data Mining Working Group
2020- Program committee for IEEE/ACM Transactions on Computational Biology and Bioinformatics
2021- Program committee for IJCAI International Joint Conference on Artificial Intelligence
2021- Program Committee for AMIA Informatics Summit
2021- Program committee for AAAI Conference on Artificial Intelligence Social Impact
2024- Program committee for Clinical NLP at NACCL
2024- Committee, AMIA Women in AMIA Pathways
2024- Program committee for Association for Computational Linguistics (ACL)
2024- Program committee for IEEE/ACM Connected Health: Applications, Systems and Engineering Technologies (CHASE)
2024- Program committee for International Conference on Artificial Intelligence in Medicine (AIME)
2024- Program committee for IEEE International Conference on Healthcare Informatics (ICHI)
2024 Co-chair of International Conference on Intelligent Biology and Medicine (ICIBM 2024), Local Committee
2024- Program committee for Conference on Empirical Methods in Natural Language Processing (EMNLP)
2024- Scientific Program Committee for AMIA Informatics Summit 2025
2024- Reviewer for International Conference on Learning Representations (ICLR)

Academic Editor
PLoS Biology
Ad-hoc journal reviewer:
Nature Communication
npj Digital Medicine
JAMA Network Open
Journal of Biomedical Informatics
Journal of American Medical Informatics
Briefing in Bioinformatics
ACM Transaction on Intelligent Systems and Technology
Cell Reports
Knowledge and Information System
Scientific Reports
Cerebral Cortex

CURRENT TEACHING RESPONSIBILITIES:
Fall 2019- BMI 5311 Foundation II: *Information Retrieval and Data Mining* (Invited lecture)
Spring 2022 BMI 7151 Seminar in Precision Medicine: *Drug Repurposing for COVID19 Using Graph Neural Network with Genetic, Mechanistic, and Epidemiological Validation* (invited lecture)
Fall 2022 BMI 7320 Topics for Artificial Intelligence in Cancer Discovery: *Graph Convolutional Networks and Lab (*invited lecture)
Spring 2023- BMI 6319 Data Analysis for Scientific Research in Biomedical Informatics
Fall 2023- BMI 5007 Methods in Health Data Science

MENTORING ACTIVITIES:
2019 Trainee:
Summer intern*:* Kai Zhang (Texas A&M)
2020 Trainee:
Postdoc: Kang Lin Hsieh, Junyong Jeong
GRA: Yan Chu, Tongtong Huang, Qian Qian
Summer intern: Xiaotian Ma (UC San Diego), Dulin Wang (Emory U.), Yaobin Ling (UW-Madison), Brandon Zhang (Rice U)
2021 Trainee:
Postdoc: Kang Lin Hsieh, Junyong Jeong, Pulakesh Upadhyaya, Yan Ding,
GRA: Yaobin Ling, Can Li, Dulin Wang, Kohong Lin, Xiaotian Ma
Summer intern: Daniel Amran (U of Houston)
2022 Trainee:
Postdoc: Kang Lin Hsieh, Junyong Jeong, Pulakesh Upadhyaya, Yan Ding
GRA: Yaobin Ling, Dulin Wang, Kohong Lin, Xiaotian Ma, Tongtong Huang
2023 Trainee:
Postdoc: Kang Lin Hsieh, Yu-Chun Hu
GRA: Yaobin Ling, Dulin Wang, Kohong Lin, Xiaotian Ma, Tongtong Huang
Visiting student: Weiran Jiang (Williams College)
Summer intern: Aaron Duong (Rice U)
2024 Trainee:
Postdoc: Ghulam Hussain, Tariq Bdair, Xinyan Zhao, Mohammad Azam Khan
GRA: Qidi Xu, Yaobin Ling, Dulin Wang, Xiaotian Ma, Yaoshun (Patrick) Chung, Kun Hee Kim (MD Anderson), Amarachi Njoku, Christina Van Hal
Visiting student: Teresa Zhou (Rice U)
Summer intern: Jeff Zhao (UT Austin), Mayuresh Dongare (U. Colorado Boulder), Juily Pachundkar (U. Colorado Boulder), Elizabeth Johnson

ONGOING GRANT SUPPORT:

R01AG084637 (MPI **Kim,** Jiang) Role: PI 07/01/2024 - 6/31/2029
Data-driven Subtypes of Alzheimer's disease progression for targeted treatment
The goal of this project is to identify subgroups of patients showing similar progression patterns across multimodal data, including cognitive function, brain atrophy, and biomarkers, by harmonizing multiple completed clinical trial data.

R01AG082721 (MPIs **Kim**, Jiang) Role: PI 09/01/2023 - 5/31/2028
Harmonizing multiple clinical trials for Alzheimer's disease to investigate differential responses to treatment via federated counterfactual learning
Drug development for treating Alzheimer’s disease has been very challenging due to the differential responses of patients to different treatments. This project will develop machine learning models to identify patient subgroups who respond differently to treatments. This will result in smaller, less expensive studies, more targeted Alzheimer’s disease clinical trials that expose fewer patients to experimental medications to which they are unlikely to respond.

U54HG012510 (PI Ohno-Machado) Role: Co-Investigator 06/01/22-03/31/26
A FAIR Bridge2AI Center
The Ethics & Trustworthy AI Core will coordinate the activities of ethics modules across Bridge2AI data generation projects.

U01CA274576 (Site PI Jiang) Role: Co-Investigator 06/2023-05/2026

Robust privacy-preserving distributed analysis platform for cancer research: addressing data bias and disparities
There is a growing recognition that bias caused by incomplete data in EHR is a significant factor contributing to the unfairness of ML/AI models in medicine, which can perpetuate and exacerbate health inequities and disparities. We will develop fair, robust machine learning models for addressing missing data in EHR, representing a crucial methodological gap.

R01AG083039 (MPIs Jiang, Tao, Bian) Role: Co-Investigator 07/2023 - 06/2026
An end-to-end informatics framework to study Multiple Chronic Conditions (MCC)'s impact on Alzheimer's disease using harmonized electronic health records
To develop ontology models and use them to guide the natural language processing (NLP) models to distill, organize, and convert MCC and relevant concepts into FHIR-accessible data.

U01AG079847 (MPIs Zhao, Jiang) Role: Co-Investigator 09/2023-08/2028
AIM-AI: an Actionable, Integrated, and Multiscale genetic map of Alzheimer's disease via deep learning
Develop a comprehensive, multiscale genetic map of Alzheimer's disease using deep learning techniques to enable targeted interventions and personalized treatment strategies

R01LM014520 (MPIs Ayday, Vaidya, Jiang)
Title: Accelerating Genomic Data Sharing and Collaborative Research with Privacy Protection
The main goal is to enhance genomic data sharing and collaborative research by implementing robust privacy protection measures, fostering innovation, and accelerating advancements in personalized medicine and healthcare.

SELECTED COMPLETED SUPPORT:
Robert Woods Johnson Foundation (MPIs, **Kim**, Jiang) Role: PI 09/15/19 - 09/14/21
Computational Phenotyping to Better Understand Obesity
This project will develop a dynamic pattern mining model to identify trends/patterns of weight gain (temporal phenotyping) that is most likely to be diabetic.

R01AG066749-03S1 (MPIs Jiang, **Kim**) Role: MPI 09/01/22 – 08/31/23
Finding combinatorial drug repositioning therapy for Alzheimer's disease and related dementias
Ethically optimize machine learning models with real-world data to improve algorithmic fairness
The project aims to develop novel machine-learning solutions to mitigate algorithmic unfairness by addressing data biases.

PUBLICATIONS
The first and/or corresponding authorship **highlighted**

A. Conference, Workshop paper or Abstract

1. **Kim, Yejin**; Park, Yong Hyun; Lee, Ji Youl; Choi, In Young; Yu, Hwanjo; Discovery of prostate specific antigen pattern to predict castration resistant prostate cancer of androgen deprivation therapy, ACM Ninth International Workshop on Data and Text Mining in Biomedical Informatics in conjunction with CIKM, Oct 23, 2015, Melbourne, Australia
2. Lee, Sun Jung; Yu, Sung Hye; **Kim, Yejin;** Choi, Inyoung; Development of Integrated Data and Prediction System Platform for the Localized Prostate Cancer, Medinfo 2019
3. **Kim, Yejin**; Giancardo; Luca; Pena Danilo; Jiang, Xiaoqian; Finding Discriminative Subgroups of Brain Regions using Tensor Factorization, ACM SIGKDD Workshop on Applied Data Science for Healthcare, 2019

4. **Kim, Yejin;** Lhatoo, Samden; Zhang Guo-Qiang; Chen, Luyao; Jiang, Xiaoqian; Temporal Phenotyping for Transitional Disease Progress: an Application to Epilepsy and Alzheimer's Disease, ACM International Conference on Web Search and Data Mining (WSDM) Health Day, 2020
5. Amran, Albert; Lin, Yaobin; **Kim, Yejin**; Bernstam, Elmer; Jiang, Xiaoqian; Schulz, Paul E; Influenza vaccination is associated with a reduced incidence of Alzheimer's disease: Epidemiology/Risk and protective factors in MCI and dementia, Alzheimer's & Dementia, 16, e041693, 2020
6. **Kim, Yejin**; Jiang, Xiaoqian; Causal pathway to analyze racial disparities in ADRD, Alzheimer's & Dementia, 16, e037081, 2020
7. Ling, Yaobin; **Kim, Yejin;** Jiang, Xiaoqian; Savitz, Sean I; Prediction of Individual Treatment Effects of Rehabilitation for Post-stroke Patients, 13th World Stroke Congress, 2020
8. **Kim, Yejin**; Chen, Luyao; Jiang, Xiaoqian; Savitz, Sean. I. Population-Based Drug Repurposing to Treat Thrombosis in COVID-19 Patients. Stroke, 2021, 52(Suppl_1), AP106-AP106.
9. **Kim, Yejin;** Jiang, Xiaoqian; Samuel, Sophie; Savitz, Sean. I. Reducing Laboratory Testing in Hospitalized Stroke Patients Using Machine Learning. Stroke, 2021, 52(Suppl_1), AP386-AP386.
10. Hsieh, Kang-lin; Jiang, Xiaoqian; **Kim, Yejin**. The Combination of Graph Neural Network and Multi-Task Learning Model for Cancer Drug Combination Prediction, AMIA, 2021 San Diego
11. Tariq, Muhammad Bilal; Ling, Yaobin; Savitz, Sean. I.; Fann, Yang. C.; Jiang, Xiaoqian; **Kim, Yejin**. Machine learning to predict high risk adverse events in treatment trials for intracerebral hemorrhage. International Stroke Conference, 2023.
12. Turnbull, Adam; **Kim, Yejin**; Zhang, Kai; Mao, Tianrui; Jiang, Xiaoqian; Mormino, Elizabeth; Henderson, Victor; He, Zihuai; Lin, F. Vankee. Mechanisms linking proteins and structural brain integrity to biological markers of Alzheimer's Disease. Alzheimer's & Dementia (AAIC) 2023
13. Lee, Junseok; Kim, Sungwon; Hyun, Dongmin; Lee, Namkyeong; **Kim, Yejin**; Park, Chanyoung Park. Deep single-cell RNA-seq data clustering with graph prototypical contrastive learning. ICML 2023 Workshop on Computational Biology.
14. Tariq, Muhammad Bilal; Ling, Yaobin; Savitz, Sean I; Jiang, Xiaoqian; **Kim, Yejin**; Machine Learning To Predict High-Risk Adverse Events In Treatment Trials For Intracerebral Hemorrhage, International Stroke Conference 2023
15. Pan, Alan P; Wozny, Joseph; Schaefer, Caroline; Nicolas, Charlie; Bako, Abdulaziz; Potter, Thomas; Caballero, Elizabeth; Nair, Rejani; **Kim, Yejin**; Jiang, Xiaoqian; Ganduglia, Cecilia; Vahidy, Farhaan S. Utilization of Sequential Pattern Mining for Evaluating Post-Acute Transitions of Care among Ischemic and Hemorrhagic Stroke Survivors: Analysis of Medicare Beneficiaries in the State of Texas, AMIA Clinical Informatics Summit, 2024
16. Xu, Qidi; **Kim, Yejin**; Schulz, Paul; Gottlieb, Assaf. Utilizing Fitness Trackers for Early Detection of Mild Cognitive Impairment: A Pilot Study on Non-Invasive Digital Biomarkers, AMIA 2024

B. Refereed Original Articles in Journals (and Full papers in Major Conference)

17. **Kim, Yejin**; Jeong, Jo-Eun; Cho, Hyun; Jung, Dong-Jin; Kwak, Minjung; Rho, Mi Jung; Yu, Hwanjo; Kim, Dai-Jin; Choi, In Young; Personality factors predicting smartphone addiction predisposition: Behavioral inhibition and activation systems, impulsivity, and self-control, PloS one, 11, 8, e0159788, 2016, Public Library of Science San Francisco, CA USA (selected as the 10% most cited paper in PloS 2016)

18. Park, Yong Hyun; **Kim, Yejin**; Yu, Hwanjo; Choi, In Young; Byun, Seok-Soo; Kwak, Cheol; Chung, Byung Ha; Lee, Hyun Moo; Kim, Choung Soo; Lee, Ji Youl; Is lymphovascular invasion a powerful predictor for biochemical recurrence in pT3 N0 prostate cancer? Results from the K-CaP database, Scientific reports, 6,1,1-7, 2016, Nature Publishing Group
19. **Kim, Yejin**; Park, Yong Hyun; Lee, Ji Youl; Choi, In Young; Yu, Hwanjo; Discovery of prostate specific antigen pattern to predict castration resistant prostate cancer of androgen deprivation therapy, BMC Medical Informatics and Decision Making, 16,1,1-9,2016, BioMed Central
20. **Kim, Yejin**; Sun, Jimeng; Yu, Hwanjo; Jiang, Xiaoqian; Federated tensor factorization for computational phenotyping, Proceedings of the 23rd ACM SIGKDD International conference on knowledge discovery and data mining,887-895, 2017
21. **Kim, Yejin**; El-Kareh, Robert; Sun, Jimeng; Yu, Hwanjo; Jiang, Xiaoqian; Discriminative and distinct phenotyping by constrained tensor factorization, Scientific reports,7,1,1-12, 2017, Nature Publishing Group
22. Choi, Jingyun; Rho, Mi Jung; **Kim, Yejin**; Yook, In Hye; Yu, Hwanjo; Kim, Dai-Jin; Choi, In Young; Smartphone dependence classification using tensor factorization, PloS one,12,6,e0177629, 2017, Public Library of Science San Francisco, CA USA
23. Choi, Jingyun; **Kim, Yejin**; Kim, Hun-Sung; Choi, In Young; Yu, Hwanjo; Tensor-Factorization-Based phenotyping using group information: Case study on the efficacy of statins, Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics, 516-525,2017,
24. **Kim, Yejin**; Choi, Jingyun; Chong, Yosep; Jiang, Xiaoqian; Yu, Hwanjo; DiagTree: Diagnostic Tree for Differential Diagnosis, Proceedings of the 2017 ACM on Conference on Information and Knowledge Management,1179-1188, 2017,
25. Choi, Jingyun; **Kim, Yejin**; Kim, Hun-Sung; Choi, In Young; Yu, Hwanjo; Phenotyping of Korean patients with better-than-expected efficacy of moderate-intensity statins using tensor factorization, Plos one,13,6,e0197518, 2018, Public Library of Science San Francisco, CA USA
26. **Kim, Yejin**; Jiang, Xiaoqian; Chen, Luyao; Li, Xiaojin; Cui, Licong; Discriminative Sleep Patterns of Alzheimer's Disease via Tensor Factorization, AMIA 2019, 2019
27. **Kim, Yejin**; Kim, Kwangseob; Park, Chanyoung; Yu, Hwanjo; Sequential and Diverse Recommendation with Long Tail., IJCAI, 2019, 2740-2746,2019
28. **Kim, Yejin**; Jiang, Xiaoqian; Giancardo, Luca; Pena, Danilo; Bukhbinder, Avram S; Amran, Albert Y; Schulz, Paul E; Multimodal phenotyping of Alzheimer's disease with longitudinal magnetic resonance imaging and cognitive function data, Scientific reports,10,1,1-10, 2020, Nature Publishing Group
29. Vance, Carroll; **Kim, Yejin**; Zhang, Guoqiang; Lhatoo, Samden; Tao, Shiqiang; Cui, Licong; Li, Xiaojin; Jiang, Xiaoqian; Learning to detect the onset of slow activity after a generalized tonic–clonic seizure, BMC Medical Informatics and Decision Making,20,12,1-8,2020, BioMed Central
30. Lee, Sun Jung; Yu, Sung Hye; **Kim, Yejin**; Kim, Jae Kwon; Hong, Jun Hyuk; Kim, Choung-Soo; Seo, Seong Il; Byun, Seok-Soo; Jeong, Chang Wook; Lee, Ji Youl; Prediction system for prostate cancer recurrence using machine learning, Applied Sciences,10,4,1333, 2020, MDPI
31. **Kim, Yejin**; Lhatoo, Samden; Zhang, Guo-Qiang; Chen, Luyao; Jiang, Xiaoqian; Temporal phenotyping for transitional disease progress: An application to epilepsy and Alzheimer's disease, Journal of biomedical informatics,107,103462, 2020, Academic Press
32. Raisaro, Jean Louis; Marino, Francesco; Troncoso-Pastoriza, Juan; Beau-Lejdstrom, Raphaelle; Bellazzi, Riccardo; Murphy, Robert; .., **Kim, Yejin**, ..Bernstam, Elmer V; Wang, Henry; Bucalo, Mauro; Chen, Yong; SCOR: A secure international informatics infrastructure to investigate COVID-19, Journal of the American Medical Informatics Association,27,11,1721-1726,2020,Oxford University Press
33. Chong, Yosep; Lee, Ji Young; **Kim, Yejin**; Choi, Jingyun; Yu, Hwanjo; Park, Gyeongsin; Cho, Mee Yon; Thakur, Nishant; A machine-learning expert-supporting system for diagnosis

prediction of lymphoid neoplasms using a probabilistic decision-tree algorithm and immunohistochemistry profile database, Journal of pathology and translational medicine,54,6,462-470, 2020, The Korean Society of Pathologists and the Korean Society for Cytopathology

34. Zhu, Cong; **Kim, Yejin**; Jiang, Xiaoqian; Lhatoo, Samden; Jaison, Hampson; Zhang, Guo-Qiang; A lightweight convolutional neural network for assessing an EEG risk marker for sudden unexpected death in epilepsy, BMC Medical Informatics and Decision Making,20,12,1-8,2020, BioMed Central
35. Mier, Juan C; **Kim, Yejin**; Jiang, Xiaoqian; Zhang, Guo-Qiang; Lhatoo, Samden; Categorisation of EEG suppression using enhanced feature extraction for SUDEP risk assessment, BMC Medical Informatics and Decision Making,20,12,1-6, 2020, BioMed Central
36. **Kim, Yejin**; Jiang, Xiaoqian; Lhatoo, Samden D; Zhang, Guo-Qiang; Tao, Shiqiang; Cui, Licong; Li, Xiaojin; Jolly, Robert D; Chen, Luyao; Phan, Michael; A community effort for automatic detection of postictal generalized EEG suppression in epilepsy, BMC Medical Informatics and Decision Making, 2020, 12,1-7,2020, BioMed Central
37. Lamichhane, Bishal; **Kim, Yejin**; Segarra, Santiago; Zhang, Guoqiang; Lhatoo, Samden; Hampson, Jaison; Jiang, Xiaoqian; Automated detection of activity onset after postictal generalized EEG suppression, BMC Medical Informatics and Decision Making,20,12,1-10, 2020, BioMed Central
38. Chong, Yosep; Thakur, Nishant; Lee, Ji Young; Hwang, Gyoyeon; Choi, Myungjin; **Kim, Yejin**; Yu, Hwanjo; Cho, Mee Yon; Diagnosis prediction of tumours of unknown origin using ImmunoGenius, a machine learning-based expert system for immunohistochemistry profile interpretation, Diagnostic pathology,16,1,1-9, 2021, BioMed Central
39. **Kim, Yejin**; Zheng, Shuyu; Tang, Jing; Jim Zheng, Wenjin; Li, Zhao; Jiang, Xiaoqian; Anticancer drug synergy prediction in understudied tissues using transfer learning, Journal of the American Medical Informatics Association,28,1,42-51, 2021, Oxford Academic
40. **Kim, Yejin**; Suescun, Jessika; Schiess, Mya C; Jiang, Xiaoqian; Computational medication regimen for Parkinson's disease using reinforcement learning, Scientific reports,11,1,1-9, 2021, Nature Publishing Group
41. Hsieh, Kanglin; Wang, Yinyin; Chen, Luyao; Zhao, Zhongming; Savitz, Sean; Jiang, Xiaoqian; Tang, Jing; **Kim, Yejin**; Drug repurposing for COVID-19 using graph neural network and harmonizing multiple evidence, Scientific reports, 11, 1, 1-13, 2021, Nature Publishing Group
42. Ding, Yan; Jiang, Xiaoqian; **Kim, Yejin**; Relational graph convolutional networks for predicting blood–brain barrier penetration of drug molecules, Bioinformatics,38,10,2826-2831, 2022, Oxford University Press
43. **Kim, Yejin**; Zhang, Kai; Savitz, Sean I; Chen, Luyao; Schulz, Paul E; Jiang, Xiaoqian; Counterfactual analysis of differential comorbidity risk factors in Alzheimer's disease and related dementias, PLOS Digital Health, 1,3,e0000018,2022, Public Library of Science San Francisco, CA USA
44. Bukhbinder, Avram S; Ling, Yaobin; Hasan, Omar; Jiang, Xiaoqian; **Kim, Yejin**; Phelps, Kamal N; Schmandt, Rosemarie E; Amran, Albert; Coburn, Ryan; Ramesh, Srivathsan; Risk of Alzheimer's Disease Following Influenza Vaccination: A Claims-Based Cohort Study Using Propensity Score Matching, Journal of Alzheimer's Disease, 1-14, 2022, IOS Press
45. Dai, Y., Yu, H., Yan, Q., Li, B., Liu, A., Liu, W., Jiang, X., **Kim, Yejin**., Guo, Y. and Zhao, Z., 2022. Drug-Target Network Study Reveals the Core Target-Protein Interactions of Various COVID-19 Treatments. Genes, 13(7), p.1210.
46. Upadhyaya, Pulakesh; Zhang, Kai; Li, Can; Jiang, Xiaoqian; **Kim, Yejin**. Scalable Causal Structure Learning: Scoping Review of Traditional and Deep Learning Algorithms and New Opportunities in Biomedicine. JMIR Med Inform. 2023 Jan 17;11:e38266. doi: 10.2196/38266. PMID: 36649070.
47. Ling, Yaobin; Upadhyaya, Pulakesh; Chen, Luyao; Jiang, Xiaoqian; **Kim, Yejin**. Emulate randomized clinical trials using heterogeneous treatment effect estimation for personalized

treatments: Methodology review and benchmark. J Biomed Inform. 2023 Jan;137:104256. doi: 10.1016/j.jbi.2022.104256. Epub 2022 Nov 28.
48. Hsieh, Kang-Lin; Plascencia-Villa, German; Lin, Ko-Hong; Perry, George; Jiang, Xiaoqian; **Kim, Yejin**. Synthesize heterogeneous biological knowledge via representation learning for Alzheimer's disease drug repurposing. iScience. 2022 Nov 26;26(1):105678. doi: 10.1016/j.isci.2022.105678.
49. Lin, Ko-Hong; Hsieh, Kang-Lin; Jiang, Xiaoqian; **Kim, Yejin**. Integrating Comorbidity Knowledge for Alzheimer's Disease Drug Repurposing using Multi-task Graph Neural Network. AMIA Informatics Summit 2023
50. Lee, Junseok; Kim, Sungwon; Hyun, Dongmin; Lee, Namkyeong; **Kim, Yejin**; and Park, Chanyoung. "Deep single-cell RNA-seq data clustering with graph prototypical contrastive learning." *Bioinformatics* 39, no. 6 (2023): btad342.
51. Bhattarai, Kritib; Rajaganapathy, Sivaraman; Das, Trisha; **Kim, Yejin**; Yongbin Chen, The Alzheimer's Disease Neuroimaging Initiative, The Australian Imaging Biomarkers and Lifestyle Flagship Study of Ageing, Qiying Dai, Xiaoyang Li, Xiaoqian Jiang, Nansu Zong, Using artificial intelligence to learn optimal regimen plan for Alzheimer's disease, Journal of the American Medical Informatics Association, 2023;, ocad135
52. Harris, Kristofer, Yaobin Ling, Avram S. Bukhbinder, Luyao Chen, Kamal N. Phelps, Gabriela Cruz, Jenna Thomas, **Kim**, **Yejin;** Xiaoqian Jiang, and Paul E. Schulz. "The Impact of Routine Vaccinations on Alzheimer's Disease Risk in Persons 65 Years and Older: A Claims-Based Cohort Study using Propensity Score Matching." *Journal of Alzheimer's Disease* Preprint: 1-16. 2023
53. Upadhyaya, Pulakesh; Ling, Yaobin; Chen, Luyao; **Kim, Yejin**; Jiang, Xiaoqian; Inferring Personalized Treatment Effect of Antihypertensives on Alzheimer's Disease Using Deep Learning, IEEE International Conference on Healthcare Informatics (ICHI), 2023.
54. Bhattarai, Kritib; Rajaganapathy, Sivaraman; Das, Trisha; **Kim, Yejin**; Chen, Yongbin; Dai, Qiying; Li, Xiaoyang; Jiang, Xiaoqian; Zong, Nansu. Learning Physician's Treatment for Alzheimer's Disease based on Electronic Health Records and Reinforcement Learning. 2023 IEEE 11th International Conference on Healthcare Informatics (ICHI)
55. Lin, Ko-Hong; Zhu, Jay-Jiguang; Smith, Judith A; **Kim, Yejin**; Jiang, Xiaoqian. An End-to-end In-Silico and In-Vitro Drug Repurposing Pipeline for Glioblastoma. 2023 IEEE 11th International Conference on Healthcare Informatics (ICHI), 738-745
56. Ling, Yaobin; Tariq, Muhammad Bilal; Tang, Kaichen; Fann, Yang C.; Aronowski, Jaroslaw; Savitz, Sean I.; Jiang, Xiaoqian; **Kim, Yejin**. An Interpretable Framework to Identify Responsive Subgroups from Clinical Trials Regarding Treatment Effects: Application to Treatment of Intracerebral Hemorrhage. PLOS Digital Health, 3(5), p.e0000493. 2024
57. Ma, Xiaotian; Shyer, Madison; Harris, Kristofer; Wang, Dulin; Hsu, Yu-Chun; Farrell, Christine; Goodwin, Nathan; Anjum, Sahar; Bukhbinder, Avi; Khan, Tanveer; Schulz, Paul E; Jiang, Xiaoqian; **Kim, Yejin**. Deep Learning to Predict Rapid Progression of Alzheimer's Disease From Pooled Clinical Trials: A Retrospective Study. PLOS Digital Health 3 (4), e0000479, 2024
58. Wang, Dulin; Ling, Yaobin; Jiang, Xiaoqian; **Kim, Yejin**. Characterizing Treatment Responders in Completed Alzheimer's Disease Clinical Trials. AMIA Annual Symposium 2024
59. Li, Tianhao; Shetty, Sandesh; Kamath, Advaith; Jaiswal, Ajay; Jiang, Xiaoqian; Ding, Ying; and **Kim, Yejin** "Cancergpt: Few-shot drug pair synergy prediction using large pre-trained language models." npj Digital Medicine (*2024)*
60. Turnbull Adam; **Kim Yejin**; Zhang K, Jiang X, He Z, Henderson VW, Lin FV. Age-associated proteins explain the role of medial temporal lobe networks in Alzheimer's disease. Geroscience. 2024 Jul 31. doi: 10.1007/s11357-024-01291-0.
61. Wang, Dulin; Ma, Xiaotian; Schulz, Paul; Jiang, Xiaoqian; **Kim, Yejin**. Clinical outcome-guided deep temporal clustering for disease progression subtyping. J Biomed Inform. 2024 Sep 30;158:104732. doi: 10.1016/j.jbi.2024.104732. PMID: 39357664.

62. Xu, Qidi; **Kim, Yejin**; Chung, Karen; Schulz, Paul; Gottlieb, Assaf. Prediction of Mild Cognitive Impairment Status: Pilot Study of Machine Learning Models Based on Longitudinal Data From Fitness Trackers. JMIR Form Res. 2024 Jul 18;8:e55575. doi: 10.2196/55575. PMID: 39024003; PMCID: PMC11294783.

C. Invited Articles (Reviews, Editorials, etc.) in Journals

63. Li, Ziyi; Jiang, Xiaoqian; Wang, Yizhuo; **Kim, Yejin;** Applied machine learning in Alzheimer's disease research: omics, imaging, and clinical data, Emerging topics in life sciences,5,6,765-777, 2021, Portland Press Ltd.

D. Other Professional Communications

I. Presentations (by local, regional, national, international)

Local

1. "Data Science and Applications", Korea University WISET, Sep 2018, Seoul, Korea
2. "Tensor Factorization for Computational Phenotyping", KCC 2017 Spotlight Session for Young Women Scholars Jun 2017, Jeju, Korea
3. "Tutorial: Matrix/Tensor factorization and its applications," Department of Medical Informatics, Catholic University of Korea, Nov 2017, Seoul, Korea
4. "DiagTree: Diagnostic Tree for Differential Diagnosis," The Korea Society of Medical Informatics, Nov 2017, Seoul, Korea
5. "Tensor Factorization for Computational Phenotyping", School of Biomedical Informatics, University of Texas Health Science Center at Houston, Nov 2018, Houston, Texas
6. "Discovering Underlying Concepts with Tensor Factorization", Center for Health Security and Phenotyping, School of Biomedical Informatics, University of Texas Health Science Center at Houston, Jan 18, 2019, Houston, Texas
7. "Secure Healthcare Machine Learning", Cizik School of Nursing, University of Texas Health Science Center at Houston, Jan 2020
8. "Scalable Causal Structure Learning and Application to Identifying Clinical Pathways from Electronic Health Records," Data Science and Informatics Core for Cancer Research, School of Biomedical Informatics, the University of Texas Health Science Center at Houston, Feb 2023, Houston, Texas
9. "Counterfactual analysis of differential comorbidity risk factors in Alzheimer's disease and related dementias," Biostatistics/Epidemiology/ Research Design (BERD) core McGovern Medical School, University of Texas Health Science Center at Houston, March 2023, Houston, Texas
10. "Predicting Personalized Treatment Effect from Data," Annual Advisory Council, School of Biomedical Informatics, University of Texas Health Science Center at Houston, March 2023, Houston, Texas, March 2023
11. "AI/ML to Identify Personalized Treatment Effect from Data" Artificial Intelligence in Biomedicine and Healthcare: A Collaborative Workshop Presented by McWilliams School of Biomedical Informatics at UTHealth Houston
12. "AI for Computational Phenotyping and Predicting Accelerated Aging" at UTHealth Houston Neuroscience Research Center and Aging Research Symposium, October 2024
13. "AI for Computational Phenotyping and Predicting Accelerated Aging" at UTHealth Houston Neuroscience Research Center, September 2024

Regional and National

14. "Data Science and Applications", Department of Management Engineering, Ulsan National Institute of Science and Technology, May 2018, Ulsan, Korea

15. “Optimized Medication Regimen for Parkinson’s Disease using Reinforcement Learning”, Rice Data Science Conference, October 2019, Houston, Texas
16. “Temporal Phenotyping for Transitional Disease Progress: an Application to Epilepsy and Alzheimer’s Disease,” ACM International Conference on Web Search and Data Mining (WSDM) Health Day, 2020, Houston, Texas
17. “Identifying Subtypes of Progression in Alzheimer’s Disease,” Rice Data Science Conference, November 2022
18. “Finding Discriminative Subgroups of Brain Regions using Tensor Factorization” ACM KDD workshop on Applied Data Science for Healthcare, August 4, 2019, Anchorage, Alaska
19. “What Large Language Models Can and Can’t do,” the Web Conference 2023 Health Day, May 2023
20. “CancerGPT: Few-shot Drug Pair Synergy Prediction using Large Pre-trained Language Models,” NIH Bridge2AI, Ethics Core, Invited talk, May 2023
21. “CancerGPT: Few-shot Drug Pair Synergy Prediction using Large Pre-trained Language Models,” PathAI, Inc, Invited talk for Research Series, August 2023
22. “Innovating Drug Development using Machine Learning,” Department of Psychiatry and Behavioral Science, SUNY Upstate, October 2023
23. “AI/ML to Identify Personalized Treatment Effects from Data,” A Collaborative Workshop Presented by McWilliams School of Biomedical Informatics at UTHealth Houston, January 2024
24. "AI for Computational Phenotyping and Predicting Accelerated Aging" at SUNY Upstate, November 2024

International

25. “DiagTree: Diagnostic Tree for Differential Diagnosis”, International Conference on Information and Management, Nov 2017, Singapore
26. “Federated Tensor Factorization for Computational Phenotyping”. The 23rd ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Aug 14, 2017, Halifax, Nova Scotia, Canada
27. “Interpretable and Federated Tensor Factorization for Computational Phenotyping”, Department of Computer Engineering, Hong Kong Baptist University, June 2018, Hong Kong
28. “Sequential and Diverse Recommendation with Long Tail”, International Joint Conference on Artificial Intelligence, Aug 14, 2019, Macau, China
29. “Causal pathway to analyze racial disparities in ADRD”, Alzheimer’s Association International Conference, July 2020, Amsterdam, Netherlands and Online
30. “Population-Based Drug Repurposing to Treat Thrombosis in COVID-19 Patients”, International Stroke Conference, 2021, Online
31. “Innovating Drug Development using Machine Learning”, Department of Industrial and System Engineering, KAIST, July 2021, South Korea
32. “Heterogeneous Treatment Effect and its Application to Healthcare”, Korea Causal Inference Summer Session, KAIST, July 2021, South Korea
33. “Innovating Drug Development using Machine Learning”, Department of Management Engineering, Ulsan National Institute of Science and Technology, July 2022, Ulsan, Korea
34. Heterogeneous Treatment Effect and its Application to Healthcare”, Korea Causal Inference Summer Session, KAIST, July 2022, South Korea
35. “Revolutionizing Healthcare with Medical Artificial Intelligence: Innovations in Diagnosis and Personalized Treatment”, Graduate School of Data Science, Seoul National University of Korea, May, 2023, South Korea

II. Video

1. Introduction to Federated Tensor Factorization, KDD 2017, Video
2. Introduction to Heterogeneous Treatment Effect Estimation and its Application to Biomedical Data, Korea Causal Inference Workshop, 2021-2022 Video
3. Stroke Buster Podcast, “Counterfactual analysis between cerebrovascular disease and Alzheimer’s disease, Institute of Stroke and Cerebrovascular Disease, UTHealth, 2022, Podcast

G. Visiting Professorships: N/A

H. Patents: N/A