YEJIN (JAYLYNN) KIM

(+82) · 10 · 2968 · 4170 ⋄ yejin.kim89@gmail.com
https://yejinjkim.github.io

EDUCATION

Pohang University of Science and Technology, Pohang, Korea *March 2012 – August 2017*
Ph.D. in Computer Science
Thesis title: Federated and Interpretable Tensor Factorization for Computational Phenotyping
GPA: 3.89/4.3

Pohang University of Science and Technology, Pohang, Korea *March 2008 – February 2012*
B.E. in Industrial Engineering
GPA: 3.83/4.3, Magna cum laude

University of California, San Diego, La Jolla, CA *September 2015 – August 2016*
Visiting Scholar in the Department of Biomedical Informatics

RESEARCH HIGHLIGHTS

I studied machine learning, data mining, matrix/tensor factorization, and recommendation system. I authored 9 papers from top-tier computer science conferences and high impact journals.

- Developing sequential recommender system of diverse and novel item by recurrent neural network, predicting next item based on previous sequence of items (Ongoing)
- Developed federated tensor factorization by alternating direction method of multiplier, overcoming privacy policy barriers (KDD 17)
- Developed novel diagnosis process framework by discrete optimization and partially observed Markov process with an application to immunohistochemistry test, reducing diagnosis time up to 6% and the number of tests up to 19% (CIKM 17)
- Developed interpretable tensor factorization by joint optimization with spectral clustering, generating data-driven knowledge with domain experts (Scientific Reports 17)

EXPERIENCES

Data Scientist in Advanced Recommendation Team, Kakao Corp.
Pangyo, Korea, *Aug 2017 – Present*
I am developing sequential recommender system by recurrent neural network (by Pytorch)

Data Scientist, Seoul St. Mary Hospital
Seoul, Korea, *July 2013 – July 2017*
Collaborating with various domain experts, I analyzed real-world data from various source and types (EMRs, clinical notes, and sensor data) using machine learning techniques (deep learning, matrix/tensor factorization, random survival forest, and discrete optimization).

- Predicted smartphone addiction from multivariate sequential sensor log using convolutional neural network and recurrent neural network (by Python and Keras)
- Derived users' behavioral patterns using tensor factorization (by Matlab and PostgreSQL)

- Predicted survival functions from EMRs with many missing values and noise using random survival forest and discrete optimization (by R and CPLEX)
- Derived frequent sequential patterns (by Java), Bayesian network, optimal threshold of new biomarkers, and various prediction models (by R)

Visiting scholar, University of California San Diego
La Jolla, California, *Sep 2015 – August 2016*
I developed federated tensor factorization models to extract phenotypes without transferring privacy-sensitive data (KDD17, Scientific Reports 17).

TECHNICAL SKILLS

Computer Languages	Python, Matlab, R (Advanced) / Java (Intermediate)
Databases	PostgreSQL (Advanced) / Microsoft SQL (experienced)
General skills	Pytorch, Keras, NumPy, scikit-learn, Linux
Languages	English, Korean

PUBLICATIONS

[9] **Kim Y**, Choi J, Chong Y, Jiang X, Yu H. DiagTree: Diagnostic Tree for Differential Diagnosis. CIKM 2017.

[8] Choi J, **Kim Y**, Kim H, Choi IY, Yu H. Tensor-Factorization-Based Phenotyping using Group Information: Case Study on the Efficacy of Statins. ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics (ACM-BCB). 2017.

[7] **Kim Y**, Sun J, Yu H, Jiang X. Federated Tensor Factorization for Computational Phenotyping. ACM SIGKDD 2017.

[6] Choi J, Rho MJ, **Kim Y**, Yook IH, Yu H, Choi IY, Kim D, Smartphone Over-dependence Classification using Tensor Factorization. PLoS One 2017.

[5] **Kim Y**, El-Kareh R, Sun J, Yu H, Jiang X. Discriminative and distinct phenotype by nonnegative tensor factorization. Scientific Reports. 2017.

[4] **Kim Y**, Park YH, Lee JY, Choi IY, Kim DH, Yu H. Discovery of Prostate Specific Antigen Pattern to Predict Castration Resistant Prostate Cancer of Androgen Deprivation Therapy. InProceedings of the ACM Ninth International Workshop on Data and Text Mining in Biomedical Informatics 2015 Oct 22 (pp. 13-13). ACM.

[3] **Kim Y**, Park YH, Lee JY, Choi IY, Yu H. Discovery of prostate specific antigen pattern to predict castration resistant prostate cancer of androgen deprivation therapy. BMC Medical Informatics and Decision Making. 2016;16(Suppl 1):63.

[2] Park YH, **Kim Y**, Yu H, Choi IY, Byun SS, Kwak C, Chung BH, Lee HM, Kim CS, Lee JY. Is lymphovascular invasion a powerful predictor for biochemical recurrence in pT3 N0 prostate cancer? Results from the K-CaP database. Scientific reports. 2016;6.

[1] **Kim Y**, Jeong JE, Cho H, Jung DJ, Kwak M, Rho MJ, Yu H, Kim DJ, Choi IY. Personality factors predicting smartphone addiction predisposition: behavioral inhibition and activation systems, impulsivity, and self-control. PLoS One. 2016 Aug 17;11(8):e0159788.