JAY J. LEE

(518) 242-9117 | W 123rd Street, NY 10027 | jl5307@columbia.edu
LinkedIn: http://www.linkedin.com/in/jay-j-lee

EDUCATION

COLUMBIA UNIVERSITY New York, NY
M.A. in Statistics, GPA: 3.5/4.0 Sep 2018 – Dec 2019
Concentration: Machine Learning and Natural Language Processing

SUNGKYUNKWAN UNIVERSITY Seoul, Korea
B.S. in Industrial Engineering, Major GPA: 3.68/4.00 March 2012 – Feb 2016
B.Econ. in International Economics and Trade

SKILLS

Programming Languages: R, Python and Java
Languages: Korean(native); English(fluent); Chinese(basic)

PROFESSIONAL EXPERIENCE

COLUMBIA UNIVERSITY DEPARTMENT OF BIOMEDICAL INFORMATICS New York
Research Associate Jan 2019 – Present

- Actively conducting researches that lies intersection of informatics and machine learning
- Use Natural Language Processing and Deep Learning methods to leverage a large amount of electronic health record (EHR)
- Support data cleanse and statistical data analysis of EHR data

SEOUL NATIONAL UNIVERSITY COLLEGE OF MEDICINE Seoul
Researcher Oct 2018 – Jan 2019

- Worked on 'Data-driven Occupational Diseases Detection System' project(funded by Ministry of Science and Technology)
- Developed automated data crawler using R and Amazon Web Services to collect data from various social network services
- Designed a document classifier using various Text Mining and Natural Language Processing techniques

REPUBLIC OF KOREA ARMY Daegu, Korea
Signal Officer (First Lieutenant) Mar 2016 – June 2018

- Administered tactical radio communication and video conference systems covering southern area of South Korea
- Analyzed error data in signal systems on a daily basis to predict future error occurrences and to forestall predicted errors
- Lead signal and communication supporting task force deployed at ROK/U.S. combined forces base

SUNGKYUNKWAN UINVERSITY DEPARTMENT OF INDUSTIRAL ENGINEERING Suwon, Korea
Research Intern Jan 2015 – August 2015

- Participated in the 'Multi-Layer Data Visualization' project, funded by the National Research Foundation
- Developed a data visualization model to assess healthcare technology and products based on U.S. patent data

RESEARCH

WORKING PAPERS

"Robust distributed representation of concepts in OMOP Common Data Model (OMOP CDM)"

The goal of this study is to generate a robust distributed representation for any concept defined in OMOP CDM and allow this representation to be used as an input for other machine learning tasks based on the OMOP CDM. I enable the representation to learn the relationships between concepts in ontology level and also in patient level using a novel method. An abstract of the research is accepted for the poster session at the "2019 Observational Health Data Science and Informatics (OHDSI) Symposium".

"Human disease recommender system using wide and deep learning"

I use wide and deep learning to learn distributed representation of terms of human diseases and apply the learned representation to develop t a recommender system for human diseases. I believe this study will uncover various latent relationships between human diseases.

PUBLISHED OR UNDER REVIEW

1. C. Liu, C. Ta, J. Rogers, Z. Li, **J. Lee**, A. Butler, N. Shang, F. Kury, L. Wang, F. Shen, L. Ena, C. Friedman, H. Liu, C. Weng. Ensembles of Natural Language Processing Systems for Portable Phenotyping Solutions. Journal of Biomedical Informatics, to appear

ACTIVITIES

OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS (OHDSI) New York
Collaborator Sep 2019 - Present

- Actively contribute to OHDSI community by participating weekly phenotyping group meeting

ACCENTURE BUSINESS CASE COMPETITION Seoul
Team Leader Jan 2016

- Developed corporate strategies for a global entertainment company to achieve successful market infiltration in East Asia

MENTORING CLUB FOR FRESHMAN Suwon
President March 2012 – Feb 2013

- Counseled 15+ student members on how to choose proper level of class and to set up course plans
- Held academic seminars on a monthly basis to help club members excel in their academics