Sulyun Lee

Iowa City, IA
sulyun-lee@uiowa.edu 319-512-5979 Personal website LinkedIn Github

EDUCATION

University of Iowa **Expected May 2022**
Ph.D., Information Science *Iowa City, IA*
- Advisor: Dr. Kang Zhao

Handong Global University **Feb. 2017**
B.S., Computer Science and Engineering *Pohang, South Korea*

RESEARCH AND PROFESSIONAL EXPERIENCE

Research Assistant **May 2019 – Present**
Department of Internal Medicine, University of Iowa *Iowa City, IA*
- Supervisor: Dr. Phillip Polgreen and Dr. Linnea Polgreen

Research Assistant **Aug. 2017 – May 2019**
College of Pharmacy, University of Iowa *Iowa City, IA*
- Supervisor: Dr. Linnea Polgreen

Research Assistant **Mar. 2016 – Dec. 2016**
School of Computer Science and Electrical Engineering, Handong Global University *Pohang, South Korea*
- Supervisor: Dr. Shin Hong, Dr. Youngsup Kim

Python Textbook Translator **June 2015 – Feb. 2017**
School of Computer Science and Electrical Engineering, Handong Global University *Pohang, South Korea*
- Translated *Introduction to Computation and Programming in Python* by John V. Guttag from English to Korean.

PROJECTS

Prediction of NFL Team Performance via Network Embedding (Codes) **Oct. 2020 – Present**
- Predicted future performance of teams in National Football League based on the hierarchical collaboration network of coaches using network embedding technique.
- Used the coaching records data collected from football statistics websites.
- Developed a supervised framework that learns team network embedding by hierarchically aggregating coach features using Graph Attention Network.
- Implemented framework with Pytorch and and Networkx.

Learning Healthcare Embedding from Hospital Interactions **Sep. 2020 – May 2021**
- Proposed a framework that learns embeddings of patients, medications, inpatient rooms, and doctors based on patients' hospital interactions to improve healthcare and to predict the future patient risks.
- Collaborated with professors and students from the Department of Computer Science, Univerity of Iowa

Team Collaboration for COVID-19 Research (Codes) **May 2020 – Mar. 2021**
- Explored the collaboration patterns of research teams in studying newly emerged field of COVID-19 via modeling the research team performance using explanatory models.
- Used COVID-19 Open Research Dataset and collected additional dataset by scraping bioRxiv/medRxiv websites and using Scopus, PubMed, and PMC APIs.
- Implemented with Python Statsmodel package for modeling and statistical analyses.
- Collaborated with professors from the College of Business, University of Iowa

Prediction of HIV/AIDS in Syringe Exchange Program Implementation (Codes) **Feb. 2020 – May 2020**
- Simulated on the number of HIV/AIDS diagnoses and deaths when the number of syringe exchange programs (SEP) is increased or the implementation of SEP is legalized in each state in the US.
- Used the state-level historical HIV/AIDS infection records provided by CDC and the SEP information in each state from The Foundation for AIDS Research (amFAR).
- Implemented with Python Sklearn package to train the prediction models, such as linear regression and random forest.
- Supported the importance of implementing more SEPs in states in significantly reducing the new diagnosis and death rates of HIV/AIDS.
- Collaborated with a doctoral student from the College of Pharmacy, University of Iowa.

- Simulations showed that increasing the number of SEPs in every state could decrease the diagnoses and deaths of HIV/AIDS by 7%.

Risk analysis for CF carriers **May 2019 – May 2021**

- Conducted the case-control study by analyzing the risks of electrolyte and fluid disorders, heart diseases, pancreatitis, diabetes, and asthma for the cystic fibrosis (CF) carriers, who inherited one copy of CF gene mutation from one of their parents.
- Used the MarketScan Research Database which includes the medical claims of inpatient visits, outpatient visits, and pharmaceutical prescriptions.
- Used logistic regression models to compute the incidence and severity of CF-related diseases for CF carriers using the glm package from R.
- Collaborated with a professor from the College of Pharmacy and a professor from the Department of Internal Medicine, University of Iowa.

Link Prediction in Online Health Community (OHC) for Smoking Cessation (Codes) **Aug. 2018 – Jan. 2020**

- Predicted the future interactions among the users of online health community for smoking cessation.
- Used the interaction records from BecomeAnEx community where users share information and emotionally support with each other through diverse communication channels.
- Considered the heterogeneity of interaction networks that comes from different types of interactions among users and used the network topological structures, network embedding, and topic modeling (LDA) of texts as predictors to predict the future interactions.
- Discovered that considering different types of communications helps 8% better prediction of future connections than treating communications homogeneously.
- Results implied that OHC recommender system can benefit from recommending users who have the similar communication activities in group discussion channels.
- Used Python Networkx, Gensim, Sklearn, and Keras packages for analyses.
- Collaborated with a professor from the College of Business and a student from Department of Computer Science.

Delay Response Analysis on Crowd Funding Platform (Codes) **May 2018 – Dec. 2020**

- Used an explanatory model (linear regression) to discover the association between backers' sentiment and the entrepreneurs' response to production delays in a crowd funding platform.
- Web scraped the comments and update contents posted by entrepreneurs and backers from Indiegogo Crowd Funding Website.
- Model explained that when the delivery of funded products was delayed, the backers are more likely to have positive sentiment if entrepreneurs showed the promise of prompt delivery.
- Used Python Statsmodel, Sklearn, TextBlob, and BeautifulSoup packages.
- Collaborated with two professors from the College of Business.

Treatment Optimization for Acute Myocardial Infarction (AMI) patients **Aug. 2017 – May 2019**

- Developed a model that optimizes the regimens of medications for AMI patients using Medicare claims database.
- Trained a neural network model that recommends the optimized regimens of AMI-related medications by maximizing the survival probability, given patients' demographics, comorbidities, and procedures.
- Estimated that the most recommended treatments should be received by 30% of AMI patients, but actually be taken by only 2%.
- Used R nnet package for neural network training.
- Collaborated with a professor from the College of Pharmacy and a professor from the College of Business, University of Iowa.

Classroom allocation (Codes) **Sep. 2016 – Dec. 2016**

- Analyzed four-year course registration data from Handong Global University to optimize the classroom allocation that minimizes the travel distance of students.
- Applied genetic algorithm.
- Used Python to implement the genetic algorithm.

Developing Education Support Webpage **Jul. 2015 – Jun. 2016**

- Developed an education support webpage for students in developing countries.
- Operated Amazon Web Server, implemented user interface by JSP and JavaScript, dynamically designed the webpage using Bootstrap, and stored user data using MySQL.
- Collaborated with a South Korean company, Big Systems Inc.

Movie Data Analysis Project **Sep. 2015 – Dec. 2015**

- Generated a machine learning model that predicts the box office hits of movies, given the directors, actors, and other movie information.
- Web scraped to collect data of 2,000 actors and 700 directors.
- Discovered that directors, actors, and distributors are the major factor that affects the box office hit of movies.

TECHNICAL SKILLS

Languages: Python, R, Java, C/C++, HTML/CSS, JavaScript, SQL, Java Server Pages (JSP)
Technologies/Frameworks: Linux, GitHub, SPSS
Python packages

- Data Management: Numpy, Pandas
- Graph/Network: Networkx, Igraph
- Machine learning: Sklearn
- Statistical analysis: Statsmodel
- Deep learning: Pytorch, Keras
- Web scraping: BeautifulSoup
- Visualization: Matplotlib, Seaborn

TEACHING EXPERIENCE

Workshop Instructor **Aug. 2019 – Present**
Iowa Social Science Research Center *Iowa City, IA*

- Python Workshops

Teaching Assistant **Aug. 2017 – Dec. 2017**
Department of Computer Science, University of Iowa *Iowa City, IA*

- Topics in Computer Science I (JavaScript)

Teaching Assistant **Mar. 2016 – Dec. 2016**
Department of Computer Science and Electrical Engineering, Handong Global University *Pohang, South Korea*

- Introduction to Big Data

Teaching Assistant **Feb. 2016**
Department of Computer Science and Electrical Engineering, Handong Global University *Pohang, South Korea*

- Python Camp

CONFERENCE AND WORKSHOP PAPERS

Sulyun Lee, Hankyu Jang, Kang Zhao, Michael S. Amato, and Amanda L. Graham. "Link Prediction in an Online Health Community for Smoking Cessation." *KDD workshop on Mining and Learning with Graphs.* Virtual Meeting, 2020.

Sulyun Lee, Hankyu Jang, Kang Zhao, Michael S. Amato, and Amanda L. Graham. "Multi-Relational Link Prediction for an Online Health Community." *INFORMS Data Science Workshop*, Seattle, WA, 2019.

PRESENTATIONS

Sulyun Lee, Kang Zhoa, Ning Li. "Understanding The Research Collaborations During COVID-19 Pandemic." *INFORMS Annual Meeting.* Virtual Meeting, 2020.

Sulyun Lee, Hankyu Jang, Kang Zhao, Michael S. Amato, and Amanda L. Graham. "Link Predictions in an Online Health Community for Smoking Cessation." *KDD Workshop On Mining and Learning with Graphs.* Virtual Meeting, 2020.

Sulyun Lee, Hankyu Jang, Kang Zhao, Michael S. Amato, and Amanda L. Graham. "Link Predictions For Social Networks in Online Health Communities." *INFORMS Annual Meeting.* Seattle, WA, 2019.

CERTIFICATES

Structuring Machine Learning Projects **Aug. 2018**
Coursera

Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization **Jul. 2018**
Coursera

Neural Networks and Deep Learning **Jun. 2018**
Coursera

Machine Learning **Mar. 2016**
Coursera

R Programming **Nov. 2015**
Coursera

edX Verified Certificate for Introduction to Computer Science and Programming Using Python **Aug. 2015**
edX

HONORS AND AWARDS

Graduate Fellowship **Sep. 2020 – Aug. 2021**
Interdisciplinary Graduate Program in Informatics

PROFESSIONAL SERVICE

Leader of Big Data Academic Society **Mar. 2016 – Dec. 2016**
Handong Global University