Sulyun Lee

Iowa City, IA

sulyun-lee@uiowa.edu 319-512-5979 https://sulyunlee.github.io/ 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*

TECHNICAL SKILLS

Languages: Python, R, Java, C/C++, HTML/CSS, JavaScript, SQL
Technologies/Frameworks: Linux, AWS, GitHub, Latex, SPSS
Python packages
- Data management: Numpy, Pandas, JSON
- Machine learning: Sklearn
- Web scraping: BeautifulSoup
- NLP: Gensim, NLTK
- Graph/Network: NetworkX, Igraph
- Statistical analysis: Statsmodel
- Visualization: Matplotlib, Seaborn
- Deep learning: Pytorch, Keras

RESEARCH INTEREST

Graph embedding, Graph neural network, Data mining in social networks, Data science, Team collaboration network

RESEARCH AND PROFESSIONAL EXPERIENCE

Graduate Research Assistant **May 2019 – May 2021**
Department of Internal Medicine, University of Iowa *Iowa City, IA*
- Supervisor: Dr. Phillip Polgreen and Dr. Linnea Polgreen

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

Undergraduate 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

Representation Learning in Hierarchical Collaboration Networks (Codes) **Oct. 2020 – Present**
Pytorch, NetworkX, Gensim
- Proposed a team embedding model that predicts the team performance based on the hierarchical collaboration network using the deep learning framework
- Achieved 9% gain of prediction performance in NFL coach dataset

Learning Dynamic Heterogeneous Embedding in Networks **Sep. 2020 – May 2021**
Pytorch
- Proposed a network embedding model from a dynamic network with heterogeneous interactions
- Achieved 35% gain in classification performance using a hospital interaction network
- Collaborated with professors and students from the Department of Computer Science, University of Iowa

Team Collaboration for COVID-19 Research (Codes) **May 2020 – Mar. 2021**
Statsmodel, Gensim, NetworkX, BeautifulSoup
- Explored the collaboration patterns that lead to team success among researchers in studying COVID-19

- Collected dataset by web-scraping and using APIs, applied an LDA topic model for representing texts, generated collaboration networks, and modeled linear/logistic regression
- Collaborated with professors from the College of Business, University of Iowa

HIV/AIDS Prediction with Syringe Exchange Program (SEP) (Codes) **Feb. 2020 – May 2020**

Sklearn

- Simulated on the number of HIV/AIDS diagnoses and deaths when the number of SEPs is increased using linear regression and random forest models.
- Discovered that increasing SEPs reduces the HIV/AIDS diagnoses and deaths by 7%
- Collaborated with a doctoral student from the College of Pharmacy, University of Iowa

Risk Analysis for Cystic Fibrosis (CF) Carriers **May 2019 – May 2021**

R – Glm and Lm packages

- Analyzed the risks of electrolyte and fluid disorders, heart diseases, pancreatitis, diabetes, and asthma for CF carriers using odds ratio analysis
- Discovered that CF carriers are at higher risks of CF-related diseases than others
- Collaborated with a professor from the College of Pharmacy and a professor from the Department of Internal Medicine, University of Iowa

Link Prediction in an Online Health Community (Codes) **Aug. 2018 – Jan. 2020**

Sklearn, Gensim, Keras, NetworkX

- Predicted the future interactions among online health community users on a multi-relational network
- Achieved 8% better prediction by implementing an LDA topic modeling, learning network embedding, and modeling Logistic Regression, Random Forest, Adaboost, and MLP
- Collaborated with a professor from the College of Business and a student from Department of Computer Science, University of Iowa

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

Statsmodel, BeautifulSoup, TextBlob

- Explored the association between backers' sentiment and entrepreneurs' delay response in a crowdfunding platform with regression analysis, and supported the importance of entrepreneurs' promises on prompt delivery
- Web-scraped a crowdfunding website, performed sentiment analysis, and modeled a linear regression
- Collaborated with two professors from the College of Business, University of Iowa

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

R – Nnet package

- Proposed a model that optimizes the medications for AMI patients by increasing survival probability up to 40%
- Trained a neural network model that predicts patient survival with a 0.8 AUC score
- Collaborated with a professor from the College of Pharmacy and a professor from the College of Business, University of Iowa

TEACHING EXPERIENCE

Graduate Teaching Assistant **Fall 2021, Fall 2017**

Department of Computer Science, University of Iowa *Iowa City, IA*

- Courses: Analyzing Data for Informatics, Topics in Computer Science I (JavaScript)

Workshop Instructor **Aug. 2019 – Present**

Iowa Social Science Research Center *Iowa City, IA*

- Workshops on Python programming

Undergraduate Teaching Assistant **Feb. 2016 – Dec. 2016**

School of Computer Science and Electrical Engineering, Handong Global University *Pohang, South Korea*

- Course: Introduction to Big Data
- 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 Zhao, 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

Session Chair **Oct. 2021**
INFORMS Annual Meeting 2021

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