Sulyun Lee

Iowa City, IA

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

RESEARCH AND PROFESSIONAL EXPERIENCE

Graduate Research Assistant **May 2019 – May 2021**

Department of Internal Medicine, University of Iowa *Iowa City, IA*

- Optimized the medical regimen and achieved an increase of patients' survival rate by 40% for heart attack patients using Neural Network models
- Implemented feature selection and parameter tuning to train Lasso Logistic Regression, Random Forest, Support Vector Machines, and Neural Networks that predict patient survival with a 0.8 AUC score

Graduate Research Assistant **Aug. 2017 – May 2019**

College of Pharmacy, University of Iowa *Iowa City, IA*

- Validated the high risks of several diseases for genetic disease carriers using regression analyses
- Provided data-driven evidence to support a hypothesis for genetic disease carriers' risks and improve public health policy

Undergraduate Research Assistant **Jun. 2015 – Feb. 2017**

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

- Optimized classroom assignments in a university by minimizing students' travel distance using a genetic algorithm
- Translated *Introduction to Computation and Programming in Python* by John V. Guttag from English to Korean

RESEARCH INTEREST

Graph embedding, Graph neural network, Data mining in social/business networks, Data science, Machine learning

TECHNICAL SKILLS

Machine Learning: Regression, Random Forest, AdaBoost, XGBoost, SVM, KNN, Clustering, PCA

Deep Learning: MLP, CNN, RNN, LSTM, GNN

Programming Language: Python, R, SQL, Java, C/C++

Technologies/Frameworks: Linux, AWS, Git, SPSS

Python packages: Numpy, Scikit-learn, PyTorch, Keras, NLTK, Gensim, NetworkX, BeautifulSoup, Matplotlib, Seaborn

EDUCATION

University of Iowa **Expected May 2022**

Ph.D., Information Science *Iowa City, IA*

Handong Global University **Feb. 2017**

B.S., Computer Science and Engineering *Pohang, South Korea*

DATA SCIENCE PROJECTS

Representation Learning in Hierarchical Collaboration Networks (Codes) **Oct. 2020 – Present**

PyTorch, Python

- Introduced a novel deep learning model that learns representations of teams with hierarchical structures and predicts team performance
- Achieved a 9% gain in predicting NFL team success from coaches' hierarchical collaboration networks

Learning Dynamic Heterogeneous Representation in Networks **Sep. 2020 – May 2021**

PyTorch, Python

- Proposed a deep learning model that learns representations of dynamic networks with heterogeneous interactions
- Achieved a 35% gain in the classification performance for patient mortality using a hospital interaction network

Team Collaboration for COVID-19 Research (Codes) **May 2020 – Mar. 2021**

Python, Regression analysis, NLP, Web scraping

- Performed regression analysis and text analysis (LDA modeling) to identify the collaborative patterns leading to scholars' team success

- Presented suggestions about knowledge transfer of scholars for a research team success to collaborators from the Department of Management and Organizations

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

Python, Scikit-learn

- Proposed the importance of SEP implementation by demonstrating the decreases of HIV/AIDS by 7% using Linear Regression and Random Forest
- Provided the data-driven insights for an important policy-making process to a collaborator from the College of Pharmacy

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

Python, Scikit-learn, Keras, NLP

- Presented a model that predicts future interactions among online health community users by distinguishing different communication types
- Achieved an 8% gain in prediction performance by implementing Logistic Regression, Random Forest, AdaBoost, and Neural Networks to train the model

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

Python, Regression analysis, Scikit-learn, Web scraping, NLP, Sentiment analysis

- Discovered the entrepreneurs' behavioral factors associated with customers' satisfaction on crowdfunding platforms using regression analysis
- Implemented a model that predicts delays in entrepreneurs' production using Random Forest, AdaBoost, and XGBoost
- Provided data-driven insights into how entrepreneurs should deal with production delays for successful businesses

TEACHING EXPERIENCE

Python Workshop Instructor **Aug. 2019 – Present**

Iowa Social Science Research Center *Iowa City, IA*

- Workshops: Data Management and Analysis, and Network Analysis with NetworkX, Introduction to Programming

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)

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 for university students

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 and Kang Zhao. "Representation Learning in Hierarchical Collaboration Networks for Team Performance Prediction." *INFORMS Annual Meeting.* Anaheim, CA, 2021

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

HONORS AND AWARDS

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

IGPI Student Travel Funding **Oct. 2019**
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