Hankyu Jang

Iowa City, IA ◇ (+1) 319-512-6129 ◇ hankyu-jang@uiowa.edu ◇ jhkmath@gmail.com
hankyujang.github.io ◇ github.com/HankyuJang ◇ linkedin.com/in/hankyujang

RESEARCH INTERESTS

Graph mining - graph neural networks (GNNs), transformer network
Social network analysis - network embedding, diffusion, link prediction, node classification
Applied machine learning - Predictive modeling, inference

TECHNICAL SKILLS

Deep learning: PyTorch, Keras
Graph-Tools: DGL, NetworkX
NLP: HuggingFace, NLTK
ML-Tools: Scikit-Learn, Numpy, Pandas, Scipy
Cloud: AWS-[EC2, S3, Athena], Terraform
Database: MySQL
Visualization: Matplotlib, igraph, Tableau
Languages: Python, R, C/C++, Shell, Java

EDUCATION

University of Iowa, Iowa City, IA — *Expected May 2023*
Ph.D. in Computer Science — GPA: 3.98 / 4.0

Indiana University, Bloomington, IN — *May 2018*
M.S. in Data Science — GPA: 3.80 / 4.0

Handong Global University, Pohang, Korea — *Aug 2016*
B.S. in Computer Science & Management, Cum Laude — GPA: 3.94 / 4.5

WORK EXPERIENCE

Machine Learning and Data Science Intern — *May 2021 - Aug 2021*
American Family Insurance, Madison, WI, USA
- Developed graph attention base model that detects suspicious data entry and further suggests the correct entry on multiple data columns on AmFam auto collision claims dataset
- Designed data validation model for Homesite insurance that corrects data using claim description

Graduate Research Assistant — *Jan 2019* - Current
Dept. of Computer Science, University of Iowa, Iowa City, IA, USA
Advisor: Dr. Alberto Segre and Dr. Sriram Pemmaraju
- Developed deep learning architecture that learns node embeddings in heterogeneous, dynamic graphs
- Designed data-driven models to detect asymptomatic infections
- Developed agent-based disease simulators (COVID-19 simulator)

Intern — *Mar 2016 - Jun 2016*
Dept. of Life Science, Handong Global University, Pohang, Korea
- Built machine learning classifiers to classify protein sequences (Mentor: Dr. Ah-ram Kim)

Research Assistant — *Mar 2015 - Feb 2016*
Dept. of Computer Science, Handong Global University, Pohang, Korea
- Organized Python programming teaching materials (Mentor: Dr. Youngsup Kim)
- Translated "*Introduction to Computation and Programming in Python*" from English to Korean

TEACHING EXPERIENCE

Graduate Teaching Assistant, *Dept. of Computer Science, University of Iowa, Iowa City, IA, USA*
Courses: Discrete Structures *Aug 2018 - Dec 2018*

Python Camp Instructor, *Dept. of Computer Science, Handong Global University, Pohang, Korea*
Lectured programming at a Python camp to students from 3 universities *Feb 2016*

Teaching Assistant, *Dept. of Computer Science, Handong Global University, Pohang, Korea*
Courses: Data Structures, Java Programming *Mar 2015 - Dec 2015*

PUBLICATIONS

H. Jang, P. M. Polgreen, A. M. Segre, S. V. Pemmaraju, *"COVID-19 modeling and non-pharmaceutical interventions in an outpatient dialysis unit,"* PLoS computational biology, 2021 Paper Data Code

D.M.H. Hasan, A. Rohwer, ***H. Jang***, T. Herman, P. M. Polgreen, D. K. Sewell, B. Adhikari, S. V. Pemmaraju, *"Modeling and Evaluation of Clustering Patient Care into Bubbles"* IEEE International Conference in Healthcare Informatics (ICHI) 2021 Paper

H. Jang, P. M. Polgreen, A. M. Segre, D. K. Sewell, S. V. Pemmaraju, *"A Data-driven Approach to Identifying Asymptomatic C. diff Cases,"* In Proceedings of the ACM SIGKDD Workshop on Epidemiology meets Data Mining and Knowledge Discovery (epiDAMIK), Aug 24, 2020 Paper

S. Lee, ***H. Jang***, K. Zhao, M. Amato and A. Graham, *"Link Predictions in an Online Health Community for Smoking Cessation,"* In Proceedings of the 15th International Workshop on Mining and Learning with Graphs (MLG), Aug 24, 2020 Paper

S. Lee, ***H. Jang***, K. Zhao, M. Amato and A. Graham, *"Multi-Relational Link Prediction for an Online Health Community,"* INFORMS Workshop on Data Science, Seattle, WA, USA, Oct 19, 2019 Paper

H. Jang, S. Justice, P. M. Polgreen, A. M. Segre, D. K. Sewell, and S. V. Pemmaraju, *"Evaluating Architectural Changes to Alter Pathogen Dynamics in a Dialysis Unit,"* Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Vancouver, BC, Canada, Aug 27 - 30, 2019 **[Best Paper Award]** Paper

POSTER & DATA PUBLICATIONS

M Suneja, T. Herman, P. M. Polgreen, D. K. Sewell, S. Justice, and ***H. Jang***, "*Healthcare Personnel Movement Data.*" Kaggle, 2020, doi: 10.34740/KAGGLE/DSV/1397235 Data

J. Liang, ***H. Jang***, DMH. Hasan, P. Polgreen, S. Pemmaraju, A. Segre. Using Data Collected from a Commercial Sensor System to Inform Mathematical Models of Healthcare-Associated Infections. Infection Control & Hospital Epidemiology. Cambridge University Press. Nov 2020 Abstract Poster

HONORS & AWARDS

Best Paper Awards
H. Jang, S. Justice, P. M. Polgreen, A. M. Segre, D. K. Sewell, and S. V. Pemmaraju, *"Evaluating Architectural Changes to Alter Pathogen Dynamics in a Dialysis Unit,"* ASONAM 2019 Paper

GSS/Graduate College Presentation Travel Funds *Oct 2019*

Data Analysis Winner at 2017 Indiana Medicaid Data Challenge *Oct 2017*
- Visualized imbalance in capacity and demand of mental health treatment in the Indiana state
- Published the solution at the in.gov - Solution Visualization Presentation

PRESENTATIONS

[Invited Talk] Introduction to network science
School of Global Entrepreneurship and Information, Communication Technology (ICT), Handong Global University, May 31, 2021 - Presentation

[Workshop] A Data-driven Approach to Identifying Asymptomatic C. diff Cases
ACM SIGKDD Workshop on Epidemiology meets Data Mining and Knowledge Discovery (epiDAMIK), virtual presentation due to COVID-19, Aug 2020 - Presentation Poster

[Workshop] Multi-Relational Link Prediction for an Online Health Community
INFORMS Workshop on Data Science, Seattle, WA, USA, Oct 2019 - Presentation

[Conference] Evaluating Architectural Changes to Alter Pathogen Dynamics in a Dialysis Unit
International Symposium on Network Enabled Health Informatics, Biomedicine and Bioinformatics (HI-BI-BI). ASONAM, Vancouver, BC, Canada, Aug 2019 - Presentation

PROFESSIONAL SERVICE

Journal Reviewer
Social Network Analysis and Mining (SNAM)

Conference Reviewer and Program Committee Member
ACM KDD workshop Epidemiology meets Data Mining and Knowledge discovery (epiDAMIK)

DATA SCIENCE PROJECTS

Image Captioning using Deep Learning Pdf Poster Code
- Implemented encoder-decoder scheme from scratch to train deep learning model to caption images
- Encoder: Applied transfer learning using pretrained weights ResNet50 to encode images
- Decoder: Used LSTM network to learn mapping between image embedding and training captions

Dog Breed Classification Code
- Achieved 79% accuracy for classifying 8K dog images into 133 categories via transfer learning
- The model performance improved by 315% compared to simple CNN architectures

IMDB Movie Reviews Sentiment Classification Code
- Achieved 86% accuracy for classifying 50K IMDB movie reviews into positive or negative using MLP

Predict Daily Bike Rental Ridership Code Data
- Accurately predicted hourly bike rental counts for 10 days using MLP for regression

Kaggle Competition: Statoil/C-CORE Iceberg Classifier Challenge Pdf Code Data
- Achieved 90% accuracy for classifying satellite images into iceberg or ship using CNN
- Explored performance of KNN, Random Forests, and SVM on PCA dimension reduced data

Identification and Localization of Siren Signals Pdf Code
- Developed a framework that accurately detects presence of ambulance siren in noisy signals
- Mapped the signals to lower dimension using NMF, then trained SVM on the representation

Single Cell Classification Pdf Code Data
- Achieved 96% accuracy for classifying 3K mouse brain cells into 9 categories using SVM
- Applied PCA to reduce feature dimension from 5K to 50 without loss of model accuracy