

SANGPIL YOUM

812-830-2413 | philz0918@gmail.com | <https://github.com/philz0918>

EDUCATION

University of Florida, Gainesville, FL

Mar 2022 - present

Ph D. in Computer Science GPA : 3.83/4.00

Indiana University, Bloomington, IN

May 2021

Master of science in Data Science GPA : 3.88/4.00

SungKyunKwan University, Seoul, South Korea

Feb 2019

Bachelor in Data Science, Library and Information Science GPA : 3.92/4.50 , 4.08/4.50

University at Buffalo, Buffalo, NY

Jan 2016 - May 2016

Exchange Student GPA : 3.93/4.00

PUBLICATIONS

Journal articles

- Jang, S. H., **Youm, S.**, & Yi, Y. J. Anti-Asian Discourse in Quora: Comparison of Before and During the COVID-19 Pandemic with Machine- and Deep-Learning Approaches *Race and Justice*. 2023; 13(1): 55–79. <https://doi.org/10.1177/21533687221134690>
- Jang S. H., Gerend MA, **Youm S.**, Yi Y. J. Understanding coronavirus disease 2019 (COVID-19) vaccine hesitancy: Evidence from the community-driven knowledge site Quora. *DIGITAL HEALTH*. 2022;8. <https://doi.org/10.1177/20552076221145426>
- Kim J, **Youm S.**, Shan Y, Kim J. Analysis of Fire Accident Factors on Construction Sites Using Web Crawling and Deep Learning Approach. *Sustainability*. 2021; 13(21):11694. <https://doi.org/10.3390/su132111694>

Manuscript under review

- **Youm, S.**, Mather, B., Jayaweera, C., Prada, J., Dorr, B. First Come First Assign (FCFA): Explainable, Divergence-Aware SRL Projection
- **Youm, S.**, Jang, S. H., Choi, J. Y., Yi, Y. J. Comparison of perceptions in the US and South Korea regarding COVID-19 vaccines on community-driven knowledge sites
- Kang, P., Ko, Y. J., **Youm, S.**, Davis, S., Connaughton, D. Consumers' value perception toward sport Non-Fungible Token (NFT): Application of Artificial Intelligence (AI) analytics on Twitter Big Data

WORKING PAPERS

Development of prediction model for construction site accident through LSTM

| beautifulsoup, selenium, nltk, keras

- Extracted frequency for cause of fire accident from article and represented in time series
- Build prediction model by using LSTM(Long short-term memory)

PRESENTATIONS

Academic Conference

- **Sangpil Youm**, Sou hyun Jang, Yongjeong Yi. Understanding COVID-19 Vaccine Hesitancy: Evidence from Quora. *In proceeding of Society of Behaviour Medicine*
- Sou hyun Jang, **Sangpil Youm**, Jaeyoung Choi, Yongjeong Yi. November 2021. Comparison of Perception about COVID-19 Vaccine on Social Q & A sites : Quora(U.S.) versus Knowledge-iN(South korea) *Korean Association of Health and Medical Sociology*

TECHNICAL SKILLS

Languages and Script : Python, C, C++, Julia, R, SQL, JavaScript, HTML, CSS, XML, Erlang

Machine Learning: Classification, Linear Regression, Logistic Regression Ridge and Lasso Regression, Support Vector machine, Hidden Markov Model, K-means Clustering, Deep Learning - CNN, LSTM, BERT, Transformers

Database: Postgres, MongoDB, MySQL

Libraries: Pandas, Scikit-learn, NLTK, spaC, NetworkX, Numpy, Tensorflow, Pytorch, Keras, Selenium, BeautifulSoup, HTMLParser, requests, matplotlib, UMAP, genism, scipy, spacy

OS: Windows, Mac, Linux

Last Edited Oct, 2023

WORK EXPERIENCE

Instructor

(expected) Jan 2024 Spring

Department of Computer and Information Science and Engineering at University of Florida

- Natural Language Processing (CAI 6307)

Graduate Teaching Assistant

Jan 2023 – Present

Department of Computer and Information Science and Engineering at University of Florida

- Natural Language Processing (CAP 4641)
- Developed class projects

Graduate Research Assistant

Mar 2022 – Present

Department of Computer and Information Science and Engineering at University of Florida

- Developed generalizable algorithm for the Semantic Role Labeling (SRL) task in low/mid-resource languages
- Established *Natural Language Processing (NLP)* lab environment

Research Volunteer

June 2021 – Feb 2022

Luddy School of Informatics, Computing and Engineering at Indiana University

- Verified whether clustering method finds or not communities in case of perturbed network
- Construct method for finding out optimized modularity within constrained number of groups by using simulated annealing, developed in Julia and Python

Graduate Researcher

Oct 2020 – May 2021

Wisee

- Built and compared the performance between Deep learning model (BERT, LSTM) and machine learning model (Hidden Markov Model) for Name Entity Recognition.
- Processed Twitter data by applying NLP techniques with Spacy and NLTK and compared the efficiency between them

COURSEWORK

Programming and Computer Science: Introduction to computer science, Computer Architecture Principles, Programming Language, Advanced Data structure, Algorithms, Advanced Database System, Operating System, Data Mining, Deep Learning System, Designing Database Systems

Mathematical and Statistical Foundation: Calculus 2/3, Probability and Statistics, Linear algebra, Time series analysis, Statistical Reasoning

Data Science and Application: Natural Language Processing, Computational Social Science, Network Science, Big Data Analytics, Social Data Analytics, Introduction to Data Science, Data Science and Linguistics, Health Informatics, Information Analysis and Evaluation, Semantic Web system

LEADERSHIP, AWARDS AND OTHERS

Excellent Project Award in Python Programming Bootcamp

August 2017

SungKyunKwan University(SKKU), Seoul, South Korea

Outstanding Performance Scholarship Awards

April, September 2017

SungKyunKwan University(SKKU), Seoul, South Korea

President of the student council

December 2016 – December 2017

Department of Data Science and Library and Information Science

Army Sergeant Discharge in Full Time

December 2011 - September 2013

Republic of Korea Army, Pocheon, South Korea

REFERENCES

Available upon request