Sangpil Youm

812-830-2413 | youms@ufl.edu | https://github.com/philz0918

Education

University of Florida, Gainesville, FL Mar 2022 - present
Ph D. in Computer Science GPA : 3.88/4.00

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

Peer-Reviewed Articles

- Jayaweera, C., **Youm, S.**, Dorr, B. AMREx: AMR for Explainable Fact Verification, EMNLP Workshop FEVER 2024
- Zhang, D., Xiao, B., Gao, C., **Youm, S.**, Dorr, B. Modeling Bilingual Sentence Processing: Evaluating RNN and Transformer Architectures for Cross-Language Structural Priming, EMNLP Workshop MRL 2024
- Martinez, M., Schmer-Galunder, s., Liu, Z., **Youm, S.**, Jayaweera, C., Dorr, B. Balancing Transparency and Accuracy: A Comparative Analysis of Rule-Based and Deep Learning Models in Political Bias Classification, EMNLP Workshop SICon 2024
- **Youm, S**., Mather, B., Jayaweera, C., Prada, J., Dorr, B. DAHRS: Divergence-Aware Hallucination-Remediated SRL Projection. *The 29th International Conference on Natural Language & Information Systems (NLDB)*. 2024. https://doi.org/10.1007/978-3-031-70239-6_29
- Koo, S., **Youm, S.**, Shin, J. Cycle-GAN-based synthetic sonar image generation for improved underwater classification, Ocean Sensing and Monitoring XVI
- 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.**, Han, C., Jang, S. H., Dorr, B. "I Still Need Your Help": Online Information Seeking Behavior Among International Students on Reddit, submitted to CHI 2025.
- **Youm, S.**, Yoo, N., Jang, S. H., Dorr, B. How Mixed Emotions on Reddit Predict Depressive Symptoms in Asian American Families, submitted to Journal of Computational Social Science.
- **Youm, S.**, Jang, S.H., Kwak, H., Choi, J., Yi, Y.J. Cross-Cultural Vaccine Perceptions: Comparing COVID-19 Vaccine Public Discourse in the U.S. and South Korea using Natural Language Processing, submitted to Journal of Health and Social Behavior.

Working Papers

Social Bias Detection model for Cultural Divergence

- Develop social bias detection model for cultural divergence to avoid miscommunication

RAG for QA: Does linguistic information helps?

- Develop RAG system with POS/NER/SRL for QA tasks.

DETQUS: Decomposition-Enhanced Transformers for QUery-focused Summarization

- Develop query-focused summarization model

Presentations

Academic Conference

- **Youm, S.**, Han, C.E., Jang, S.H. July 2024. Information-seeking Behavior Among International Students on Reddi: A Comparative Analysis Before and After Start of COVID-19. *10th International Conference on Computational Social Science (IC2S2)*
- **Youm, S.**, Yoo, N., Jang, S.H. July 2024. Beyond Positive or Negative: How Mixed Emotions on Reddit Predict Depressive Symptoms in Asian American Families. *10th International Conference on Computational Social Science(IC2S2)*
- **Youm, S**, Jang, S. H., Yi, Y.J. April 2022. Understanding COVID-19 Vaccine Hesitancy: Evidence from Quora. *In proceeding of Society of Behaviour Medicine*

Work Experience

Instructor Spring 2024

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 and operated 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 & Culture (NLP&C)* 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

Wissee

- 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

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

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