Sanya Bathla Taneja
Email: sbt12@pitt.edu
Website: sanyabt.github.io
LinkedIn | GitHub | Semantic Scholar

EXPERIENCE

Graduate Student Researcher | ***February 2020 – Present***
University of Pittsburgh, Intelligent Systems Program

- Led knowledge graph construction and analysis of 700K nodes and 7M edges, combining literature-based discovery and biomedical ontologies for generation of mechanistic hypotheses for natural product-drug interactions and adverse events. Presented related work at 4 conferences, with 1 Best Poster Award and 1 first-author peer-reviewed publication (*in review*).
- Developed custom OMOP vocabulary and OBO representations for 700 natural products and constituents for pharmacovigilance and knowledge representation.
- Responsible for longitudinal EHR data extraction, data analysis, and technical development of machine learning and case-control epidemiological analyses OMOP Common Data Model for Alzheimer's disease risk factors using OHDSI methods in R and Python, with 1 peer-reviewed publication (*in review*) and 2 conference presentations.

Research Assistant | ***September 2018 – February 2020***
University of Pittsburgh, School of Medicine

- Developed natural language processing and machine learning pipelines for twitter surveillance of vaping.
- Responsible for RITHM software framework maintenance, documentation, and upkeep of the GitHub repository for real-time Twitter data mining, with 3 peer reviewed publications.

Software Development Engineer Intern | ***February – July 2018***
Amazon India

- Developed backend APIs for the Seller and Retail website using Java, Spring MVC, Coral, JavaScript, and Handlebars. Involved in adding order cancellation details to the Seller dashboard to supplement the seller website.

EDUCATION

PhD Intelligent Systems | University of Pittsburgh | 2020-2024
MS Intelligent Systems | University of Pittsburgh | 2020
Thesis*: Bayesian Networks for Diagnosing Childhood Malaria in Malawi*
B.Tech. Computer Science and Engineering | Indira Gandhi Delhi Technical University | 2018

SUMMARY

PhD candidate in Intelligent Systems and computer scientist with research experience in natural language processing, machine learning, and knowledge representation and their applications in healthcare. Proficient in Python and SQL, with strong communication and writing skills.

SKILLS AND INTERESTS

Skills and Interests: Machine Learning, Natural Language Processing, OMOP Common Data Model, EHR Data Analysis, Bayesian Networks, Knowledge Graphs, Biomedical Ontologies, Scientific Writing
Technologies: Python, SQL, R, Git, C++, RDF, OWL
Libraries: NLTK, Spacy, Pandas, Scikit-learn, Keras, Networkx, Tensorflow

PROFESSIONAL ACTIVITIES

- Student Editorial Board Member, Journal of the American Medical Informatics Association (JAMIA) | *2022-2023*
- Peer Review (BMC Bioinformatics, Intelligent Systems for Molecular Biology, AMIA Informatics Summit)
- Co-organized & coordinated discussion group on Symbolic AI and Knowledge Graphs with 20 participants and 6 guest speakers | *2022*