Seohee Sunny Yoon

akheyn029@gmail.com | 470-838-2401 | linkedin.com/in/seohee-yoon-268482263/ | https://syoon029.github.io/

Skills

- **Programming Languages**: Python, SQL, PySpark, Scala, Java, C, JavaScript, R
- **Tools**: Tableau, Pandas, Numpy, Jupyter Notebook, Scikit-learn, Excel, Matplotlib, Seaborn, D3, AWS
- **Courseworks**: Database Technology**,** Machine Learning, Natural Language Processing, Data & Visual Analytics

Work Experience

Graduate Research Assistant, SocWeb Lab - Atlanta, GA **Feb 2025 – Present**
Supervisors: Dr. MunMun De Choudhury

- Contributing to interdisciplinary research leveraging computational techniques, natural language processing, and social computing to analyze digital trace data and improve mental health outcomes
- Engaging in the development of a methodology to identify how bias against mental health entities is generated by LLMs using the toxicity rabbit hole framework

Data Analytics Intern, PTKOREA - Seoul, South Korea **Jun 2024 – Aug 2024**

- Rectified about 20 customer-reported issues daily, including dashboard access errors, incorrect data sources, and data inconsistencies using Jira and Tableau
- Validated time-sensitive data in response to client-specific requirements using Excel, SQL, and Adobe Analytics, ensuring system stability and data quality
- Spearheaded testing and validation of market dashboards with cross-functional teams, driving system performance improvements and increased user engagement

Undergraduate Research Assistant, C21U Lab - Atlanta, GA
Supervisors: Dr. Jeonghyu Lee, Dr. Meryem Yilmaz Soylu **Aug 2022 – Dec 2023**

- Enhanced assessment quality by translating problem-solving approaches into data-driven solutions, applying the Depth of Knowledge (DOK) framework
- Identified key difficulty terms, improving regression model precision by 25% in prediction precision through feature engineering with TF-IDF techniques, data augmentation, and preprocessing

Education

Georgia Institute of Technology - Atlanta, GA

Master of Science in Computer Science **Expected Dec 2025**
Bachelor of Science in Computer Science **May 2024**

Projects

Implicit Emotion Classification

- Optimized emotion dataset processing, reducing training time by 20% through regular expressions and feature engineering
- Evaluated various combinations of language models and knowledge graphs to identify the optimal solution for implicit emotion classification
- Accomplished model improvements from 64% to 88% in accuracy and F1 Score by integrating BERT with a domain-specific knowledge graph

Respiratory Diagnosis Assistant

- Developed a web-based respiratory diagnosis assistant application, utilizing GRU networks to classify respiratory diseases from audio data
- Integrated Amazon S3 and MongoDB for efficient data storage and retrieval, managing large-scale audio and metadata for real-time machine learning inference
- Improved audio classification accuracy from 70% to 83% by incorporating data augmentation techniques such as noise addition, pitch shifting, and MFCC feature extraction

Database Application Project

- Implemented relational data models and normalized schemas in MySQL to support an airline management system
- Created SQL views, queries, and procedures to analyze flight operations, crew activity, and passenger behavior, enabling scenario-based analysis and reporting