

# Seohee Sunny Yoon

akheyn029@gmail.com | 470-838-2401 | Atlanta, GA,30308 | <https://syoon029.github.io/>

## Education

### Georgia Institute of Technology

Atlanta, Georgia

Master of Science in Computer Science

Aug 2024 – Expected Dec 2025

Bachelor of Science in Computer Science

Jan 2022 – May 2024

## Experience

### Graduate Research Assistant, SocWeb Lab - Atlanta, Georgia

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 outcome
- 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 Analysis Intern, PTKOREA - Seoul, South Korea

Jun 2024 – Aug 2024

- Resolved approximately 20 customer issues daily, preventing financial losses using Jira and Tableau
- Monitored data anomalies, maintaining data quality and system stability through consistent updates with Excel, SQL, and Adobe Analytics
- Collaborated with cross-functional teams to test market dashboard functionality, identifying and resolving operational issues to enhance performance and user experience

### Undergraduate Research Assistant, C21U Lab - Atlanta, Georgia

Aug 2022 – Dec 2023

Supervisors: Dr. Jeonghyun Lee, Meryem Yilmaz Soylu

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

## Skills

- **Programming Languages:** Python, SQL, PySpark, Scala, Java, C, JavaScript, R
- **Data Analysis and Visualization:** Tableau, Scikit-learn, Excel, Matplotlib, Seaborn, D3, Adobe Analytics
- **Courseworks:** Data Structures & Algorithms, Database Technologies, Machine Learning, Data & Visual Analytics

## Projects

### Implicit Emotion Classification

Aug 2024 – Dec 2024

- 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

### Sentiment Analysis of University Student Discussion on Reddit

Aug 2024 – Dec 2024

- Prepared 3.8 million Reddit posts, ensuring dataset relevance and clarity through tokenization, stop word removal, and language filtering techniques
- Developed interactive visualizations with choropleth maps and word clouds using D3 and Tableau, providing overall sentiment trends on specific topics and keyword relations

### Respiratory Diagnosis Assistant

Feb 2024 – May 2024

- 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 applying data augmentation techniques such as noise addition and pitch shifting, and extracting MFCC features for audio classification