Seohee Yoon

|  |  |
| --- | --- |
| 470-838-2401 | [akheyn029@gmail.com](mailto:akheyn029@gmail.com) | Atlanta, GA | <https://syoon029.github.io/> |  |
|  |  |

Education

|  |  |
| --- | --- |
| **Georgia Institute of Technology** | **Atlanta, Georgia** |
| Bachelor of Science in Computer Science  Master of Science in Computer Science  GPA: 3.72 | *Jan 2022- May 2024*  *Aug 2024 - Dec 2025* |

Experience

|  |  |
| --- | --- |
| **PTKOREA** | **Seoul, South Korea** |
| *Big Data Analysis and Operation Assistant Intern* | *Jun 2024 – Aug 2024* |

* Leveraging Jira and Tableau to organize customer requirements, resolve account-related issues, and manage about 1000 customer account data effectively
* Actively participating in customer meetings across various countries with team members, recording and summarizing the minutes, and enhancing the members' work efficiency
* Managing the periodic modification and mapping of databases to cater to the needs of professionals across diverse disciplines, ensuring system consistency and reducing the likelihood of errors

|  |  |
| --- | --- |
| **Data Driven Education, Georgia Institute of Technology** | **Atlanta, Georgia** |
| *Research Assistant - Data Scientist* | *Aug 2022 – Dec 2023* |

* Identified the level of difficulty for the actual assessment from the dataset based on the Depth of Knowledge (DOK) framework utilizing Python, Google Colab, and Term Frequency - Inverse Document Frequency (TF-IDF) algorithm
* Designed a dataset for training Regressor for Difficulty and Discrimination Estimation model to predict problem DOK level
* Evaluated various algorithms (Random Forest, Decision Tree, Support Vector, Linear Regression, Ridge Regression) for optimal model using scikit-learn library
* Expanded testing data in the dataset from 50 to 900 data leveraging problems from various subjects
* Reduced mean absolute error about 50% using random forest than support vector algorithms with more data

Projects

|  |  |  |
| --- | --- | --- |
| **Respiratory Diagnosis Assistant** | *Feb 2024 – May 2024* | *September 2023* |

* Developed an automated lung sound classification Web using machine learning, improving accuracy from 70% to 83% through data augmentation and a GRU model
* Managed complex data structures with MongoDB and Amazon S3, integrating with Django for database operations
* Enabled predictive diagnostics and streamlined workflow for healthcare professionals by providing instant access to relevant audio samples and patient information

|  |  |
| --- | --- |
| **Stock Market Prediction Project** | *Feb 2024 – May 2024* |

* Predict stock price changes using several machine learning algorithms such as Support Vector Regression, Linear Regression, and LSTM
* Evaluate linear regression models using K-fold cross-validation, resulting in 95% accuracy in prediction
* Visualize prediction results from various models and compared each accuracy using tableau

|  |  |
| --- | --- |
| **Scene Recognition with Deep Learning Project** | *Nov 2023* |

* Developed a convolutional neural network (SimpleNet) with 2 convolution layers which aligns with given training dataset using PyTorch
* Implemented data augmentation techniques, normalization, and regularization to improve training accuracy and validation accuracy by 30% of the network
* Enhanced pretrained resnet from pytorch's API by modifying layer of model to get a testing accuracy of 85%

|  |  |
| --- | --- |
| **Campus Discovery Service Project** | *Aug 2022 – Dec 2022* |

* Collaboratively engineered a Campus Discovery Service Application with a team of 5 within an Agile framework comprising 6 sprints
* Designed user interface and user experience for five iOS application screens, employing JavaScript and React Native
* Utilized version control systems and enhanced team collaboration by integrating and consistently using GitHub for code management and project tracking

Skills

Python, Java, MySQL, Git, NumPy, Pandas, C#, Assembly Programming, C, React, Tableau, JavaScript, React Native