

Sunwoo Kim

sunwoo604@gmail.com | <https://sunwoo604.github.io> | <https://github.com/sunwoo604>

Education

University of California San Diego – B.S. in Data Science Minor in Computer Engineering GPA:3.7 2020 – 2024

Experience

Data Scientist, UCSD Health - George Lab – La Jolla, CA Apr 2023 – Present

- Performed development and maintenance of data engineering pipeline in Azure using Databrick while improving runtime from 4 hours to 20 minutes
- Developed data visualizing tools for Azure Database using Tableau with data analysis
- ETL of research data into Azure Data warehouse and SQL relational databases with further Data analysis

Software Engineering Intern, ESTsoft Corp. – Seoul, Korea Jul 2023 – Aug 2023

- Developed real-time video streaming tool for AI Virtual Human using Python and speech generation library from Azure
- Designed a system for speech synthesis of the user text input while displaying the video

Video ML Intern, MyEdMaster LLC – Remote Feb 2023 – Aug 2023

- Developed a video processor pipeline that extracts coordinates of body parts by object detection using YOLO, and Movenet
- Established machine learning algorithm to determine if the user is exercising with proper form using PyTorch

Fullstack Web Developer Intern, Seminaut Inc. – Remote Sep 2022 – Jan 2023

- Implemented a web page for an individual post, a system that allows the organizer of the event to remove the attendees, and a drag-and-drop table for user seeding for the bracket generation using React.js and Flask.
- Led the bracket generation algorithm and backend data structure for database optimization for SQL database

AI Business Development Intern, ESTsoft Corp. – Seoul, Korea Jul 2022 – Aug 2022

- Performed different duties such as data cleaning and data collecting, business proposal development, and building business models with market research and analysis, business translations, and recruiting models

Machine Learning and UI Intern, Arthur C Clarke Center – La Jolla, CA Jan 2022 – Jun 2022

- Developed the development server for the backend Python codes and HTML and JavaScript(Aframe) codes
- Built interactive UI, real-time changing based on the user input and Machine Learning algorithm result

Projects

Enhancing Utility Pole Monitoring With Computer Vision <https://jcheung4.github.io/DSC180B-Website/>

- Developed algorithm for collecting coordinates of each unique 360 image was taken in Google Street View
- Finetuned DETR object detection model to enhance the accuracy
- Implemented overlapping detection box prevention algorithm by area calculation with the given threshold

GWAS Database Tableau Dashboard <https://biturl.top/Bne6fe>

- Multi-tab data visualization to present an overview and the distribution of data for research data analysis
- Provided different features to filter data and visually separate them by using colors and shapes.

Tritonhack 2024 Machine Learning Starter Kit <https://jcheung4.github.io/DSC180B-Website/>

- Starter project for students participating in a Hackathon event for K-12 students, intro to Numpy, Pandas, scikit-learn, etc.
- Led the project by performing version control and designing the entire project.

Skills

Skills: Data Analysis, Data Engineering, Machine Learning, Data Science, Artificial Intelligence, Software Development, Statistics, Database Management, Cloud Computing, Business Intelligence, Data Visualization

Languages: Python, C++, C, Java, JavaScript/HTML, SQL, R

Libraries: PyTorch, NumPy, Pandas, Matplotlib, scikit-learn, Flask, openCV, Dask, aiortc, PySpark, React.js, D3.js

Tools: Azure, AWS, Excel, Jupyter Notebook, Visual Studio Code, Github, Tableau, Android Studio