KARAN SUNIL

(213)-595-0884 | karansun44@gmail.com | linkedin.com/in/karans04/ | github.com/karans04

EDUCATION

University of California, San Diego   September 2016 –
B.S in **Data Science**   September 2020
GPA – **3.65** (Provost Honors), Upper Division GPA – **3.80**

TECHNICAL SKILLS

Languages – Python, C, C++, Java, Javascript, Swift, HTML
Database – MySQL, PostgreSQL
Tools/Frameworks – Pandas, Pyspark, AWS, Dask, Scikit-learn, Django, NLTK, OpenCV, NetworkX, Librosa
Visualization – Matplotlib, Seaborn, D3.js, Tableau, Highcharts

WORK EXPERIENCE

Data Science Research Assistant   June 2019 – September 2020
Scripps Institute of Oceanography

- Built an end-to-end data pipeline to generate a 3D reconstruction of a coral reef using NASA ICESat2 satellite data and Sentinel 2 imagery.
- Used DBSCAN to cluster reef photons and map them to color pixel values of Sentinel-2 satellite images.
- Predicted the depth of the entire reef using this mapping in a Linear Regression model.
- *Presented* the project at the American Geophysical Union in San Francisco in December 2019.
- **Publication** in preparation – "Coral reef bathymetry from the joint analysis of ICESat-2 laser altimetry and Sentinel-2 multi-spectral Imaging".

Data Science Tutor   January 2018 – June 2019
Halıcıoğlu Data Science Institute

- Tutored more than 200 students in the following classes – Algorithms and Data Structures; Principles of Data Science; Practice and Applications of Data Science.
- Held discussion sessions for students to clarify difficult concepts.
- Graded and created homework assignments and examinations.

Software Engineer Intern – Machine Learning   July 2018 – September 2018
Prime Focus Technologies

- Developed an OpenCV based model that generated thumbnails from a video file.
- Generated keyframes and detected faces using FFmpeg and OpenCV's LBP Cascades.
- Applied Scikit's structural similarity index to get rid of similar frames and ranked the thumbnails using a Keras based emotion detection model.
- *Presented* by the company at the International Broadcasting Convention in Amsterdam, Netherlands.

PROJECTS

Malware Category Detection (Senior Capstone Project)   January 2020 – June 2020

- Performed a code analysis of Android apps on the Google play store using knowledge graph embeddings.
- Detected the category of malware the app belonged to, based on API calls made by the app, with an F1 score of 92%.
- Featured on Data Science department website as one of the **top 5** projects.

Amazon product category predictions   November 2018

- Predicted the category of a product using Sentiment Analysis and Natural Language Processing techniques such as TF-IDF with an accuracy of 88.128%.
- Used a Support Vector Machine (SVM) to classify items by their category.
- Ranked in the **top 5%** on the **Kaggle leaderboard**.

COURSEWORK

- Database Management
- Recommender Systems and Web Mining
- Statistical Natural Language Processing (NLP)
- Artificial Intelligence (AI) and Machine Learning (ML)
- Spatial Data Science and Applications
- Data Visualization