# KEVIN BUI

📞 (818)-823-1376  ✉ kevinbui806@gmail.com  in https://www.linkedin.com/in/kevin-bui  🏠 kevinbui2.github.io

## Education

**University of California San Diego**                                    Exp. March 2024
*Bachelors of Science in Data Science; GPA: 3.75*                                    *La Jolla, CA*

## Relevant Coursework

- Principles of Data Science
- Practice and Application of Data Science
- Data Structures and Algorithms
- Theoretical Foundations of Data Science
- Econometrics

## Certifications

- AWS Certified Cloud Practitioner
- Google Data Analytics Certificate

## Experience

**Acoustical Engineering Services**                                    Jun 2021 – Sep 2022
*Data Analyst / Acoustical Engineer*                                    *Los Angeles, CA*

- Worked at Acoustical Engineering Services to help collect, clean, and analyze data for clients to make sure the acoustics are in compliance with regulation for the client's proposal.
- Created different visualizations for the data being given to help clarify if changes were needed to be made on the design or material of the building to be in compliance of regulation.
- Helped in modeling buildings from the client's design proposal using AutoCAD to test the acoustics of the proposed building.
- Used various acoustic engineering software, such as Traffic Noise Model (TNM), Enhanced Acoustic Simulator for Engineers (EASE), and SoundPLAN to create simulation models for further analysis.

## Projects

**Song Lyrics Generator Using Language Models** | *NLTK, GPT-2, Scikit-learn, Genius API*                **June 2022**

- Created a Ngram Language Model and utilized GPT-2 to generate new song lyrics given a corpus of all of an artist's song lyrics.
- Utilized the Genius API to scrape all of the artist's songs lyrics.
- Presented a song's given highest TF-IDF and visualized the lyrics with a Word Cloud visualization.

**Spotify User's Music Statistics** | *JavaScript, Spotify Web API*                **June 2022**

- Created a website called "A Glimpse of Your Spotify Statistics" which allowed users to view their top tracks, artists, and genres in 3 different time periods.
- Utilized the Spotify Web API to gather the user's data and created data visualizations for the user's statistics, such as a pie chart to view the user's different genres.

**Youtube Comments Spam Detection** | *Youtube API, Scikit-learn*                **July 2022**

- Trained a Machine Learning model using a Naive Bayes Classifier to predict whether a comment is spam or not by use of a dataset of comments labeled spam or not-spam.
- Extracted comments from different Youtubers using the Youtube API and labeled whether the comment was spam or not to create training dataset.
- Evaluated the model to view its accuracy, precision, and recall.

**LA County Covid-19 Dashboard** | *Tableau*                **July 2022**

- Extracted datasets from *data.lacity.org* to create an interactive Covid-19 Dashboard using LA County data.
- Presented different visualizations to have users easily understand information such as the number positive cases and deaths that has occurred from the start of the pandemic and vaccination rates by neighborhood and demographic.

## Technical Skills

**Languages**: Python, SQL, R, Java
**Technologies**: AWS, Tableau, MySQL, Excel
**Other**: NumPy, Pandas, Scikit-learn, Matplotlib, Seaborn, Git