Cristobal Zamorano Astudillo

cristobalza.com | cristobal.zamorano@berkeley.edu | Berkeley, CA 94704 | +1 (310) 980-9686

EDUCATION

- **University of California, Berkeley** — Berkeley, CA
  B.A. of Computer Science and Data Science — *Expected Graduation: Spring, 2021*
  - **Major Coursework**: Machine Learning, Data Mining and Analytics, Probability & Discrete Mathematics, Databases, Data Structures, Advanced Algorithms
- **Santa Monica College** — Los Angeles, CA
  A.A/A.S. of General Science, Computer Science. Transfer requirements for Computer Science — *Aug. 2015 – Aug. 2018*
  - Dean's List, NSF Community College Scholarship

EXPERIENCE

- **Board of Directors of a Non-profit Organization** — Jan 2016 - August 2018
  University Cooperative Housing Association (UCHA) — *Los Angeles, CA*
  - **Vice-president of the Board of Directors**: Make policy and oversee the long-term UCHA's stability financially and politically. In addition, I also made weekly and quarterly financial reports of UCHA, I was responsible for the maintenance of the official files, including minutes of the Board. As Vice-president, I had to oversee and lead the overall condition of the non-profit organization along with the Executive Director and the rest of the Board members.

- **Bruins-In-Genomics (B.I.G.) Summer Undergraduate Researcher** — June 2018 - August 2018
  University of California, Los Angeles — *Los Angeles, CA*
  - **Undergraduate Summer Research Intern**: Summer Internship that combined bioinformatics and genomics. Assist with research including analyzing multiple-phenotype data using QIIME2 via command line to study gut bacteria in mice. Analyzed data sets for diseases using Python to create a database of drug interactions. Data gathering & literature research on gene-mapping analysis data visualization using Seaborn to create figures for academic papers.

PROJECTS

- **Toxic Comments Detector App - Multilabel Classifier** *(Full Stack App, Machine Learning, NLP)*
  - **Goal**: Detect toxic comments in a multi-label fashion in order to reduce the inadvertent bias of the model.
  - Dataset obtained from a Kaggle Competition. Data is composed of a large number of Wikipedia comments which have been labeled by human raters for toxic behavior.
  - [Click to Web App Link](#) - [Click to Code Repo](#)

- **Future Salary App** *(Deep Learning, Full Stack App, Machine Learning, NLP)*
  - **Goal**: Predict salary of a job given only the description/summary of a job.
  - Data obtained from over 20,000+ jobs from over 20 cities in the United States.
  - Full stack project that predicts Data related Jobs' salaries in a multi-class fashion. Each class it is a salary range.
  - [Click to Web App Link](#) - [Click to Code Repo](#)

- **Spotify Projects** *(APIs, Python, Machine Learning, Software Engineering)*
  - **PCA Construction**: Using my one-year Streaming History of songs as data, I constructed my own interactive PCA in terms of Spotify's metric features and liked songs. [Click here to see interactive PCA](#)
  - **JSON to API to CSV**: After requesting JSON file from Spotify's website, you can process every song through my code and obtain its Spotify's unique metrics.

SKILLS

- **Programming Languages**: Python (numpy, Jupyter/iPython, Pandas, Scikit-Learn, Tensorflow, Keras, GeoPandas, Seaborn, Plotly, Matplotlib, Flask), Java (Swing/JavaFX), SQL, HTML, CSS (Bootstrap 4), Javascript

- **Data Science & Miscellaneous**: A/B Testing, ETL, Data Science Pipeline (cleaning, wrangling, visualization, modeling, interpretation), Statistics, Time Series, *Regression* (Linear, Multiple-Linear, Random Forest, SGD), *Classification* (K-NN, SVM Families, Random Forest, Logistic), *Clustering* (K-Means, Spectral), *Deep Learning* (Monte Carlo Simulation, Convolutional Neural Networks) Git, NLTK, Heroku, LaTeX, Spanish (Native)