# Cristobal Zamorano Astudillo

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

## Education

- **University of California, Berkeley** — Berkeley, CA
  *B.S. of Computer and Data Science - Applied Mathematics & Modeling* — *2018 - 2021*
  - **Major Coursework**: Advanced Machine Learning, Data Mining, Probability Theory, Computer Security, Databases Design, Data Structures, Computer Architecture, Principles & Techniques of Data Science, Algorithms

## Skills

- **Languages**: Python, SQL, R, Java, C, C++, GoLang, Kotlin, HTML, CSS, Javascript
- **Technologies**: AWS, Docker, Heroku, Keras, Tensorflow, Pytorch, Flask, Spark, Git, Maven

## Highlighted Experience & Projects

- **University of California, Berkeley — Latinx Research Center** — Berkeley, CA
  - **Data Scientist Research Intern**: Jan 2021 - Present
    - Working with Professor Angela Marino constructing a data mapping demographics of minorities communities that were most impacted by the bills winners and losers of the U.S. Elections in 2020 in the state of Florida.
    - Goal is to create a database that stores key words dictionary of the most important words used to proposed bills during the Elections 2020 using NLP.
- **Future Salary App — https://future-salary.herokuapp.com/** *Python*
  - Designed a salary predictor that only uses job descriptions in order to output salary ranges
  - Implemented a word2vec Neural Network to rank text data scrapped from 20K+ different job descriptions.
  - Implemented my own Neural Network Architecture to benchmark other ML models. My Neural Network had the highest Accuracy metric of all the ML models with 85% on the test data.
- **Toxic Comments Detector App — https://toxic-text-detector.herokuapp.com** *Python*
  - Designed a multi-label classifier that reduce the inadvertent bias of a ML model to detect toxic comments.
  - Implemented an optimal feature engineering for vectorization of the text data and hypertunning parameter analysis.
  - After the model benchmark results, the final model had an overall of 90% F1 score against all toxic features.
- **Santiago's Car Accidents Predictor & Visualizations — http://www.cristobalza.com/** *Python*
  - Analyzed the most common causes of car accidents in the city of Santiago, Chile using Police data and Geodata.
  - Implemented and trained an ML model that could predict accidents upon the street and month of a year. Using a hypertunned Random Forest Regression, I was able to predict deathly accidents with an average accuracy of 83%
  - Designed a policy proposal using my work to the Chilean Transportation Ministery to prevent number of casualties
- **Gitlet** *Java*
  - Designed and implemented version control system in Java with Git-like functionalities

## Related Experience

- **University of California, Berkeley — EECS Department** — Berkeley, CA
  - **Foundations & Techniques of Data Science Lab Assistant**: Jan 2020 - March 2020
    - Co-taught weekly 2-hour lab for for DATA 100 Principles & Techniques of Data Science course taught in Python
    - Topics such as the Data Lifecycle, Machine Learning Ensemble Modeling, Data Visualization Techniques.
  - **Data Structures & Algorithms Lab Assistant**: Aug 2019 - March 2020
    - Assisted students in labs with assignments and projects for CS 61B Data Structures course taught in Java.
    - Helped students to understand the main concepts behind data structures and programming methodologies.
- **Santa Monica College — STEM Program** — Santa Monica, CA
  - **Linear Algebra and Multivariable Calculus Supplemental Instructor**: Sept 2016 - Dec 2017
    - Assisted college students with college level math classes by providing bi-weekly sessions
    - Helped students to learn course concepts, prepare for exams, and learn effective study skills.