Last Updated: 03/2021

Cristobal Zamorano Astudillo

Berkeley, CA 94704 | U.S. Citizen | +1 (310) 980-9686
cristobal.zamorano@berkeley.edu | cristobalza.com | linkedin.com/in/cristobalza/

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
    - Designing a data mapping tool to map demographics of minority communities that were impacted by the 2020 elections with Professor Angela Marino and the rest of researchers team.
    - Implement NLP techniques and methodologies on the bills that lost and won the Elections in Florida.
    - Attending weekly meetings to communicate progress, discuss challenges and troubleshooting of the project in the overall process with Professor and fellow researchers.
- **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 reduces the inadvertent bias of a Machine Learning 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.
- **Spotify Music Insights Project — http://www.cristobalza.com/** *Python*
  - Implemented the Spotify's API to obtain my music metrics to create an insight on my data.
  - Designed and implemented my own PCA algorithm using Eigenvectors from my own music data.
  - Designed and implemented my own ML recommender model that could compete against my "Discover Weekly" playlist made by Spotify.
- **Santiago's Car Accidents Classifier & 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 an ML model that could predict accidents based 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

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 in Python
    - Helped students to learn course concepts about Data Lifecycle, Machine Learning Ensemble Modeling, Data Visualization Techniques, prepare for exams, and learn effective study skills.
  - **Data Structures & Algorithms Lab Assistant**: Aug 2019 - March 2020
    - Helped students to prepare for exams, and learn effective study skills, and understand the main concepts behind data structures and programming methodologies.