Brandon Lee Concepcion

+1 (650)-892-4233 | brandon_concepcion@berkeley.edu | linkedin.com/in/brandonconcepcion | brandonconcepcion.com
github.com/brandonconcepcion

EDUCATION

University of California, Berkeley — *May 2026*
Majors: B.A. Data Science, B.A. Computer Science — *GPA: 3.81/4.0*
Technical Skills/Coursework: *Deep Learning, Neural Networks, Computer Vision, Natural Language Processing, AB Testing, Machine Learning Theory, Data Structures, Advanced Calculus, Linear Algebra, Discrete Mathematics, Statistics and Probability*
Organizations: *Data Science Society at Berkeley, Data C8 (Foundations of Data Science) Undergraduate Course Staff*
Honors/Awards: *Genentech Futurelab Scholar, 6th Annual Datathon For Social Good: 2nd Place*

SKILLS AND INTERESTS

Languages: Python, Java, SQL, RegEx, LaTeX, HTML | Learning Ruby, Javascript
Tools: Pandas, NumPy, TensorFlow, PyTorch, CV2, SciPy, StatsModels, Sci-kit Learn, Seaborn, Plotly, MatplotLib, Tableau, Streamlit
Skills: Exploratory Data Analysis, Git, Data Visualization, Classification, Clustering, Linear and Logistic Regression, Data Analytics
Soft Skills: Efficient Communication, Adaptability & Flexibility, Teamwork and Collaboration, Organizational Ability, Simple and Creative Problem Solving, Leadership, Critical-Thinking, Taking Initiative, Self-Starter, Attention to Detail, Cross-Functional
Personal Interests: Movies, Volleyball, Photography, Basketball, Gym, Swimming, Dogs, Road Trips, Music, Marvel Studios

WORK EXPERIENCE

UC Berkeley — Berkeley, CA
Data Science Lead Instructor — *Jun 2024 - Present*
- Developed and delivered educational content to a student population of **3,000+**, providing comprehensive and effective instruction through office hours, **3** review sessions, and answering of **400+** student questions
- Restructured *data8.org/su24* using **HTML**, **CSS**, and **Javascript** for front-end web development, adding **19** dynamic tabs to organize **250+** past exam problems by their scope in the course
- Led my students to achieve average exam scores in the **90th** percentile, as well as cumulative course grades in the **93rd** percentile, highest among all student instructors

Doctors Without Borders — Remote
Data Scientist Contractor — *Aug 2024 - Dec 2024*
- Utilized **Python**, **SQL** and the Armed Conflict Location & Event Data (ACLED) **API** to achieve **93% accuracy** in classifying global regions likely to experience fatalities from escalating political events, aiding in the identification of high-risk zones
- Preprocessed **2,000,000+** political events across **74** features and **180+ countries** to train two **Scikit-learn** neural networks, achieving an R^2 score of **76%** in predicting the number of fatalities for regions with escalating political events
- Developed a **Streamlit** app to visualize conflict severity and fatality predictions, enhancing humanitarian safety management

Data Science Society at Berkeley — Berkeley, CA
Instructor, Vice President — *Aug 2023 - Dec 2024*
- Managing **16** Teaching Assistants and **12** tutors to operate the "An Introduction to Real World Data Science" course, promoting accessibility by hosting educational workshops from industry and academic leaders for the **70+** diverse students in the course
- Created the course website *dssdecal.org* by utilizing **Jekyll**, **Ruby**, and **Github Actions**, and currently developing a **12**-chapter introductory data science textbook at *dssdecal.org/textbook*

University of Washington — Remote
Data Scientist Researcher — *Jan 2024 - Jun 2024*
- Coded a **Variational Auto-Encoder** (VAE) neural network in **PyTorch** and **OpenCV** to convert numerical retinal data into generative AI video simulations of retinal movement afflicted by one of three different diseases
- Implemented a data preprocessing pipeline that converts .avi files into sets of **300** individual frames
- Ran training data through a **Long-Short Term Memory** (LSTM) network to encode data into latent space, then decoded data using a **Gated Recurrent Unit** (GRU), producing video simulations in **512x512** resolution

PROJECTS

Spam Email Classifier 🔗 | *Python, Pandas, Principal Component Analysis* — *Nov 2023*
- Used **Pandas**, **NumPy**, and **RegEx** to develop an **87%** accurate classification model to predict spam emails for the School of Pharmacy, utilizing a dataset of over **7,500** points and achieving an Area Under the ROC Curve of **91%**
- Applied **Principal Component Analysis** (PCA) to reduce dimensionality by **70%**, and enhanced model performance by **5%** using GridSearch cross validation across **4** hyperparameters