JULIE LAI

(224)343-2446 ⋄ julielai@berkeley.edu ⋄ https://github.com/juliejlai/data-science

EDUCATION

University of California, Berkeley

Masters in Information and Data Science	August 2020 - August 2021
Bachelors of Arts in Physics and Astrophysics	August 2016 - May 2020

WORK EXPERIENCE

Data Scientist Intern — May 2021 - August 2021
Springbuk, Inc. — *Indianapolis, Indiana*

· Improved upon recall and precision with new machine learning model in comparison to current model
· Implemented semi-supervised machine learning models in Python to predict risk of Opioid Use Disorder using insurance health claim data
· Queried PostgreSQL databases, joining, and cleaning data for predictive modeling

Data Scientist Intern — May 2019 - June 2020
Finkbeiner Lab, Gladstone Institute of Neurological Disease — *San Francisco, California*

· Refined semi-supervised machine learning models with Python to predict ALS using patient genomic data
· Visualized trends in patient genomic data using matplotlib and seaborn graphs
· Compacted large amounts of patient genomic data into a concise dataframe of patient versus variants

PROJECTS

Global Music Recommender — August 2021

· Created a global music recommender model that utilizes KMeans clustering and validating scores using a cosine similarity metric
· Visualized song features using Javascript, CSS, and HTML to create audio-reactive art
· Connected recommender model with visualizer using Vue.js,Spotify Web API, Web Playback SDK and deployed website on Heroku

Formality Neural Machine Translation — April 2021

· Implemented a word-level LSTM sequence-to-seqeunce Neural Machine Translation models in Keras Tensorflow to translate formality using the Grammarly's Yahoo Answers Formality Corpus
· Evaluated results using a Text Formality Classifier and visualizing formality using LIME

Predicting Alzheimer's — November 2020

· Used a Kaggle Dataset of MRI Segmentation Images to classify different stages of Alzheimer's dementia
· Classified images using a variety of models including Keras, Naive Bayes, and SVC
· Filtered the data through dimension reduction

LEADERSHIP

Media Director — January 2019 - May 2020
The STEMinist Chronicles — *Berkeley, California*

· Promoting representation of women in STEM by executing the photoessay pipeline: interviewing, transcribing, editing, and publishing photoessays on social media platforms

STRENGTHS

Technical	Python (numpy, pandas, scikit-learn, matplotlib, seaborn, Altair), R, SQL, JavaScript (D3, Vega-lite), Tableau, Excel, LaTeX, Bash, LabView
Graphic Editing	Adobe Photoshop, Adobe Lightroom, Adobe Premiere Pro, Adobe Illustrator, Procreate
Languages	Conversational Chinese and Spanish