

JACKY LU

DATA SCIENTIST

✉ green.jackylu@gmail.com
☎ (408) 250-1509
📍 Cupertino, CA 95014
in jackylu
🌐 jackylu

Skills

MACHINE LEARNING

Regression
Classification
Natural Language
Processing

PROGRAMMING

Python
SQL

LIBRARIES

Pandas
NumPy
BeautifulSoup
NLTK
Spacy
Gensim
Scikit-learn

DATA VISUALIZATION

Matplotlib
Seaborn
Tableau

Education

UC San Diego
2012 to 2016
B.S. Molecular
Biology
Magna Cum Laude

Projects

r/Medicine Topic Modeling & Sentiment Analysis

- Analyzed the top monthly posts of r/medicine from January to July 2020
- Performed topic modeling to determine how the Coronavirus pandemic has shifted discourse in this subreddit
- Utilized VADER sentiment analysis to track how post sentiment has changed over time

Animal Shelter Adoption Returns Classification

- Utilize Sonoma County Animal Shelter Intake and Outake data for the past 7 years to figure out what traits are more associated with adoption returns
- Compared the following supervised classification models including: logistic regression, decision tree, random forest, Naive Bayes, and XGBoost

New York MTA Analysis - Sports Edition

- Analyzed New York MTA turnstile data to compare station activity on Yankee game days versus non-game days to determine which stations would be best to advertise in for Yankee sports fans

Craigslist Used Car Price Predictions

- Scraped SF Bay Area Craigslist car postings using BeautifulSoup
- Predicted post car prices using a linear regression and cross-validation from features such as mileage, age of the car, and post location

Employment

Metis

Data Scientist

June 2020 to Sept. 2020

- Completed a rigorous and accredited 12 week data science bootcamp

Quartzly

Life Science Product Specialist

Jan. 2018 to Apr. 2020

- Own and maintain life science product data for Quartzly's customers, suppliers, and internal teams
- Handle technical customer inquiries related to life science products including finding scientifically equivalent alternatives to discontinued or backordered items, troubleshooting quality issues, and cleaning incorrect product data for Quartzly's vast portfolio of > 10 million products
- Handle all vendor related inquiries related to price discrepancies, shipping logistics, and other fulfillment issues
- Lead on Categories & Attributes project: wrote SQL query to determine all products sold in Quartzly's fulfillment center and customer's top 5000 requested products, spearheaded effort to categorize and collect attributes for every product
- Lead on Crossmatch File project: evaluate crossmatch files from supplier partners and provide feedback to vendors on their crossmatches and how they could improve their matches in the future

Bell Biosystems

Research Associate, Cell Biology

Jan. 2017 to Dec. 2017

- Culture mammalian mesenchymal stem and T cells for use with Magnelle MRI contrast agent
- Perform immunocytochemistry and microscopy experiments on Magnelle-labeled cells
- Manage orders database and work with CFO to track laboratory expenses
- Host Bell Biosystems' company booth for SynBioBeta, a synthetic biology biotech conference