José Kabir Snell

949-212-5464 | kabirsnell@gmail.com | linkedin.com/in/kabirsnell | Portfolio: kabirjs.github.io

EDUCATION

University of California Irvine Irvine, CA
Master of Data Science *Sep. 2023 – Dec. 2024*

University of California Santa Barbara Santa Barbara, CA
Bachelor of Science in Statistics and Data Science *Sep. 2021 – June 2023*

WORK EXPERIENCE

Statistical Research Programmer Oct. 2023 – Present
Chapman University School of Pharmacy *Irvine, CA*

- Building a machine learning prediction framework using Python and R that can generate a likelihood score (0-100%) for persistent opioid use and trajectories of opioid use at the later stage of adolescence and young adulthood based on early opioid prescription history
- Using a relational database to effectively parse through over 1.5 million data points
- Performing various statistical tests and employing model validation methods to ensure that any statistical analysis included in our publications is statistically rigorous

Data Science Intern Oct. 2022 – Jan. 2023
Inogen Inc. *Goleta, CA*

- Drove informed decision-making regarding user experience by creating data visualization dashboards using Tableau and SurveyMonkey
- Pulled thousands of data-points from an AWS database using SQL queries
- Actively participated in a team that employed the agile methodology

Marketing Analytics Intern June 2022 – Aug. 2022
Automobile Club of Southern California *Costa Mesa, CA*

- Built machine learning models using R in order to improve the accuracy of forecasts by 20% for the number of memberships and leads
- Used forecasts to derive predictions for other key performance indicators and used Tableau and Excel to create an easily readable dashboard to present them
- Explained the machine learning methodology to business partners from various backgrounds in weekly meetings
- Collaborated with Media Planning team to better understand market trends and desired forecasting parameters

PROJECTS

UCSB Capstone Project | *R, Python, Keras, Tensorflow* Jan. 2023 - June 2023

- Worked on a corporate partner project in collaboration with Inogen and UCSB which leveraged Python as well, as optical character recognition libraries Keras and Tesseract, in order to automate the data extraction process from medical prescription forms
- Created a data visualization dashboard in Excel to automatically perform an analysis on the data that was retrieved using the automated prescription form retrieval
- Created a poster in order to share findings with business partners and was voted best poster at the annual UCSB Data Science capstone showcase

Music Genre Classification | *R, Python, Tensorflow, Neural Network, Random Forest* Dec. 2022

- Extracted Spotify data using an API; data set contained over 10,000 songs (data points)
- Created easy-to-read data visualizations of the many predictors vs. the response variable (genre) to better understand the data
- Utilized a decision tree, boosted forest, random forest, k-nearest neighbors, and neural network model to evaluate which classification model garnered the highest predictive accuracy
- Evaluated the best performing model on a test set of data by assessing various accuracy metrics

TECHNICAL SKILLS

Languages: R, Python, C++, SQL, SAS
Developer Tools: Git, R Studio, VS Code, Visual Studio, PyCharm, Eclipse
Libraries: Pandas, NumPy, Matplotlib, Tensorflow, Keras