José Kabir Snell

949-212-5464 | kabirsnell@gmail.com | linkedin.com/in/kabirsnell | 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*

- Extracted and cleaned a massive data set exceeding 350 million rows from the California Controlled Substance Utilization Review and Evaluation System-Prescription Drug Monitoring Program (CURES) utilizing a combination of Python, R, and SQL
- Duplicated a widely used risk score model originally trained on the CURES data set to assess its fairness and scrutinize any potential inadvertent bias against low-income individuals and underrepresented minorities
- Constructed a robust machine learning framework in Python and R, enabling the generation of likelihood scores for persistent opioid use and trajectories of opioid consumption during the later stages of young adulthood

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

- Empowered data-driven decision-making for enhancing user experience by developing interactive data visualization dashboards utilizing Tableau in conjunction with SurveyMonkey surveys
- Retrieved data sets, comprising thousands of data points, from an AWS database by executing SQL queries
- Worked within an agile-driven team environment, contributing to iterative development processes

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

- Developed advanced machine learning models using R to elevate forecast accuracy by 20% in predicting membership and lead numbers
- Leveraged forecasts to generate predictions for key performance indicators, and then harnessed the power of Tableau and Excel to craft user-friendly dashboards for clear data presentation
- Effectively communicated the machine learning methodology to diverse business partners with varying backgrounds during weekly meetings

Projects

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

- Collaborated on a corporate partnership project with Inogen and UCSB, employing Python alongside optical character recognition libraries such as Keras and Tesseract: This project aimed to automate the extraction of data from medical prescription forms, streamlining the process
- Achieved an impressive accuracy rate of 80-90% in accurately transcribing handwritten information from prescription forms
- Designed and presented an impactful poster to share research findings with business partners, earning recognition at the annual UCSB Data Science capstone showcase.

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

- Utilized the Spotify API to extract data from a data set comprising more than 10,000 songs, each representing a distinct data point
- Evaluated various classification models, including decision trees, boosted forests, random forests, k-nearest neighbors, and neural networks, to determine which model achieved the highest predictive accuracy
- Assessed the performance of the top-performing model on a separate test data set, utilizing a range of accuracy metrics to measure its predictive capabilities

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