

Ryan Voda

818-370-1304 | rvoda@g.ucla.edu | [linkedin.com/in/ryan-voda](https://www.linkedin.com/in/ryan-voda) | ryanvoda.github.io

EDUCATION

University of California, Los Angeles

Sep. 2020 – Jun. 2023

Bachelor of Science in Mathematics; GPA: 3.45

Los Angeles, CA

- Coursework: Python with Applications, Linear Algebra, Introduction to Probability, Algorithms, Optimization, Introduction to Computer Science, Real Analysis, Applied Numerical Methods

EXPERIENCE

Data Management Trainee

Aug. 2023 – Present

Sony Pictures Entertainment

Culver City, CA

- Developed a Python script to clean and manipulate data inputs, streamlining the ingestion process.
- Performed MLOps duties in SQL to build a reliable pipeline for deployment and maintenance of models.
- Identified data discrepancies and failures in Apache Airflow for major YouTube channels.

Data Management Intern

Jun. 2023 – Aug. 2023

Sony Pictures Entertainment

Culver City, CA

- Created ETL pipeline to integrate data from new vendor into Snowflake and automate workflow.
- Ingested and analyzed big data from several social media platforms, generated financial reports using BigQuery and Excel, and presented overview with key metrics and visualizations to stakeholders.
- Collaborated with data science team to make model-based features and assisted with creating machine learning model using Keras in Python.
- Maintained accessibility of databases and increased query efficiency by more than 300 percent.

Research Assistant

Mar. 2022 – Jun. 2022

(UCLA) California Health Survey Analysis with Prof. Harlin Lee

Los Angeles, CA

- Cleaned and interpreted CHIS-2020 data in Python using pandas and numpy.
- Found correlations between distress levels and 13 factors related to COVID-19 with linear regression.
- Verified 92 percent correlation by applying scipy statistical tests such as spearmanr and chisquare.
- Produced data analysis summary using plotly and matplotlib.

PROJECTS

Milk Grade Classification

Aug. 2022

- Imported dataset from <https://www.kaggle.com/datasets/prudhvignv/milk-grading> to predict milk quality using categorical and quantitative characteristics of milk.
- Cross-validated feature selection tests to find the strongest predictive features of milk grade.
- Conducted hyperparameter tuning to apply results on several appropriate Machine Learning models.

Fantasy Soccer Team Generator

Jul. 2022

- Scraped and cleaned data from <https://fplform.com/fpl-predicted-points> to obtain player statistics.
- Optimized team for maximum projected points when given constraints and game rules via Linear Programming.
- Implemented a feature that updates the team each week and makes appropriate transfers.
- Achieved a top 50 percent performance throughout the duration of the project.

Penguins Species Predictor

Feb. 2022

- Built optimal Machine Learning models for predicting penguins species from the Palmer Penguins data set.
- Performed data analysis and visualization to narrow down our selection of features.
- Designed a function to automatically clean and prepare data for train/test split.
- Evaluated performance of several models, such as Random Forest, Logistic Regression, and Clustering.

TECHNICAL SKILLS

Languages: Python, SQL, R, MATLAB, JavaScript, C++

Technologies: Excel, Tableau, Periscope, BigQuery, Snowflake

Other: Pandas, NumPy, Matplotlib, Git, Keras, TensorFlow, PyTorch, Scikit-learn