Roland Chin

BA - Data Science | minor - Computer Science

Senior at UC Berkeley graduating May '21, seeking internship / entry level job to apply & grow my skills.

249 La Honda Dr.
Milpitas, CA 95035

rond24933chn@berkeley.edu
(408)-908-9187
rchinnn.github.io

MAIN COURSES

- Principles & Techniques of Data Science
- Probability for Data Science
- Data, Inference, and Decisions
- Artificial Intelligence
- Time Series Analysis
- Applied Econometrics
- Discrete Mathematics & Probability Theory
- Computational Models of Cognition
- Human Contexts & Ethics of Data

INTERNSHIPS

zHealthEHR - Data Science Intern

May 2020 - Aug 2020

- Cleaned and queried healthcare data from chiropractors with MySQL and Excel
- Built multivariable regression model to increase optimum billing for various CPT codes, based on selected features using Scipy and Sklearn in Python

Grandmark International - Data Consulting Committee

Sept 2019 - Dec 2019

- Created a new stocking policy based on demand forecasting and customer segmentation
- Utilized regression model and confidence intervals to predict optimal turnover quantities for warehouse inventories in Python

IPMD, Inc. - Startup Intern

Aug 2018 - Jan 2019

- Collaborated with a team of Berkeley graduates to help develop an AI platform to automatically call 911 given video footage
- Gathered and labeled emergency footage data for ResNet50 and Xception models

SKILLS

Technical

- Data Cleaning
- Exploratory Data Analysis
- Visualization
- Feature Engineering
- Classification
- Regression
- Hypothesis Testing
- Time Series Analysis

Jupyter Notebook

- NumPy
- Pandas
- Matplotlib
- Seaborn
- SciPy
- Scikit-learn
- Tidyverse

Languages / Software

- Python
- R
- MySQL
- Excel
- Tableau

TEACHING

Statistics Undergrad Student Association: Education Committee

Sept 2018 - Dec 2020

- Created material for student run course on introductory data analysis
- Taught weekly programming workshops and led creation of assignments

Academic Intern for Data 8: Foundations of Data Science

Jan 2018 - Dec 2019

- Tutored students for labs, homeworks, and projects
- Assisted Berkeley MOOCLab with EdX online course creation and tested prior to launch

---

PROJECTS

Mafic Lava Flows Project

- Classified basalt rocks to examine sea floor trends and identify traits of unknown basalts
- Utilized Sklearn K-Nearest Neighbors with min-max scaling; achieved 86% accuracy

Digit Classification & Language Identification

- Implemented perceptron linear algorithm and nonlinear neural network to classify handwritten digits from MNIST as well as identify languages of words
- Found optimal hidden layer sizes, batch sizes, learning rate, and epochs

Pacman Project

- Implemented DFS, BFS, A star, and Q learning search agents for multiagent reinforcement learning Pacman game in Python
- Applied reflex agents, minimax pruning, Alpha-Beta pruning, and expectimax algorithm

Mindfulness & Executive Functions Research Study

- Created and conducted survey to gather data with SRS and cluster sampling
- Analyzed with T-test to determine effect of mindfulness meditation on cognitive performance

Spam vs Ham Email Classification

- Performed feature engineering, logistic regression, multiple linear regression, and cross validation to classify email bodies as spam
- 95% accuracy on Kaggle competition

Genre Text Classification

- Implemented K-Nearest Neighbors from scratch using Euclidean distance of frequency of words
- Predicted genre of movies and books from their texts without built in modules

Corona Borealis Galaxy Investigation

- Tested whether velocities of galaxies followed a multimodal distribution for evidence of voids and superclusters existence
- Modeled test statistic as a function of KDE bandwidth parameter