Nader Esmael

Moreno Valley, CA 92555 | 951-288-8825 | ncesmael@gmail.com | linkedin.com/in/ncesmael/ | laternader.github.io/

Education

General Assembly, Data Science Immersive Program *December 2020*

California State Polytechnic, University (Pomona), B.S. Applied Mathematics & Statistics *June 2018*

Volunteer: President & Founder of Sportsfest Club (2016-17), Athletics Chair of Barkada (2016-17), Internal Affairs Director of PASK (2015-2016), Datafest - High Ranking Participant (2018)

Coursework: Applied Probability Theory, Applied Statistics, Combinatorics, Mathematical Programming, Applied Regression, Applied Survival Analysis, Random Processes, Linear Algebra, Numerical Analysis, Differential Equations

Online: Data Science Nanodegree (June 2020), Completed SQL Bootcamp

Skills

Languages:

Proficient: Python, SQL, R

Familiar: Github, Matlab, C++, HTML

Other Relevant Skills:, Numpy, Pandas, Scikit-learn, Matplotlib, Seaborn, MS Office (Excel, Access), Tensorflow, Tableau, Cloud-Based Services (AWS, BigQuery), Web Scraping, Data Warehousing, Statistical Modeling, ggplot2, git

Machine Learning: Logistic Regression/Classification, Decision Trees/Random Forests, XGBoost

Professional Experience

Mergers & Acquisitions Research Analyst, Harvey & Company LLC, Newport Beach, CA *March 2019 - January 2020*

- Increased top-line revenue 150% by developing strategic search queries and research tactics for over 10 platforms
- Performed market research and due diligence to identify targets for acquisition based on market conditions, competitor research, industry standards
- Successfully influenced and managed resources across various workgroups and functions including assisting other analysts and interns with research development
- Designed detailed marketing materials and finance reports for bi-weekly meetings with executives and management teams to discuss research efforts and other applicable information

Conference & Events Student Assistant, Cal Poly Pomona Associated Student, Inc., Pomona, CA *August 2017 – August 2018*

- Supervised and trained six other team members including safety training, operation protocols, and time management
- Supported senior leadership through identifying process improvements and positioning the business for growth potential
- Performed assessments and benchmarks of other events and conference programs to improve schedules, delegation, and coordination

Projects

Reci-Please | Hackathon/Group Project

- Worked alongside UX/UI designers, software engineers, and data scientists to create a elderly, user-friendly web-application to find recipes based on filter algorithms and recommendation models
- Extracted, cleaned, and transformed over 500,000 recipes scraped from Food.com
- Use Python expressions to remove irrelevant characters and convert large strings into concise lists for back-end development readability, feature engineer binary columns for various holidays and meal times, convert columns to numerical data types for any future analysis and turned pandas dataframes into JSON files to be stored in a database for the web-application

Pitch Classification and Baseball | Solo Concept Project

- Achieved about 65% accuracy evaluating predicted pitch accuracy (ball vs. strike) compared to umpire calls; used a Random Forest classifier, player's pitch history, and game mechanics, beating the baseline model of 45%
- Utilized pybaseball to collect five seasons (4 million rows) of pitch data and organized it into a workable dataframe and feature engineered pitchers by determining a pitcher's repertoire percentage usage
- Determined the importance of speed, rotations, pitch start and end locations, and pitch type in predicting an umpire's pitch call

Depression or Shower Thoughts: that is the question | Solo Concept Project

- Pulled 20,000 Reddit posts using Pushift API and designed an NLP-based Random Forest classifier that differentiates if a post is a random thought or exhibits signs of depression
- Investigated r/depression and r/showerthoughts subreddits to understand the connection between language and behavior, focusing on its significance in detecting signs of depression
- Utilized the model's coefficient to determine commonly ignored keywords that lead to the misclassification of depression and achieve error margins as low as 9%