

Mona Kalika

Data Scientist

347-357-7342 | monakalika1@gmail.com | github.com/monak12 | linkedin.com/in/monakalika |

Data aficionado with a knack for economics and all things numbers. Searching for opportunities to combine my love of finding patterns and trends in the finance industry and in data science!

Skills

Python, SQL, Excel, Rstudio, Word, C++, GIT, Google Cloud
Statistical Analysis, Project Management, Time Management
Communications, Creativity, Team Leadership & Collaboration

Languages

English
Intermediate German
Conversational Spanish

Data Science Projects

Female Clothing Size Predictor and Web-App

- Generated a Multiclass Classification Model and Web-App that accurately predicts female clothing size using age, weight, and height as inputs.
- Implemented several models, including Logistic Regression, K-Nearest Neighbors (KNN), Random Forest, Extra Tree, as well as, boosting models such as AdaBoost, Gradient Boost, and a Voting Classifier.
- Produced scores 30% more accurate than the baseline accuracy score.
- Created an interactive Web-App predicting user clothing size, given user inputs.
- Pickled a Logistic Regression model to be quickly referenced and used in the Python Streamlit tool for creating a Web-App.

World Hunger: Global Crop Yield Predictor

12/2021

- Worked in a group of 4 data scientists to implement machine learning models to help the UN reach its Zero-Hunger goal by 2030.
- Used 7 models including: Multiple Linear Regression, KNN Regressor, Decision Tree, Bagging, Random Forest, Ada-Boost, Gradient Boost, and two sets of features to find the best conditions for predicting the global crop yield produced by 9 crop varieties.
- Generated models with up to 96% accuracy in predicting global crop yield.
- Performed Exploratory Data Analysis to produce violin, line, and bar plots that depict the distribution of crop yield by area harvested, over time, and by continent using matplotlib, seaborn, and geopandas.

NLP Subreddit Predictor

11/2021

- Built an NLP model that would accurately predict whether text belonged to one subreddit over another.
- Utilized Pushshift API to search and gather 3000+ individual submissions from Reddit's Relationship Advice and Jokes subreddits.
- Preprocessed data using CountVectorizer, Tokenization, Lemmatization, and Stemming to quantify the use of certain words for tuning measures.
- Created a pipeline with best params for a Logistic Regression Model and a Random Forest Model to generate accurate text predictions up to 95%. Created Confusion Matrices and generated specificity and sensitivity scores to see not only how good the model was at predicting one subreddit over the other, but rather where the model was predicting incorrectly.

Work Experience

Internal Audit Intern - I.A. Department, Touro College

02/2021 -05/2021

- Reviewed 5+ department's policies and procedure documents to ensure compliance with State and Federal laws serving 35 schools internationally.
- Audited 3 completed audits to access accuracy, verify information, and evaluate the analysis process.
- Conducted a mini audit of Touro's Purchasing Department with the guidances of my mentor, producing a Risk Control Matrix (RCM), Entry and Exist presentations, Announcement Memos, and Lead Sheets using excel and resulting in the discovery of several oversights, now up for review to improve internal controls.

Teaching Assistant - John Jay College, CUNY

11/2020 - 12/2020

SOC 324: Advanced Social Statistics

- Provided personalized one-on-one and group tutoring sessions, weekly, covering topics such as T-Tests, Chi-square tests, Correlation, ANOVA, etc. 12+ weeks.
- Reviewed assignments and provided feedback to 23 students, weekly
- Improved student understanding of Microsoft Excel/Google Sheets by demonstrating complex statistical functions; stdev, SE, confidence intervals, etc.
- Assisted students in brainstorming research questions for final projects using quantitative and qualitative socio-economic data to create descriptive statistics tables, info-graphics, regression and multi-regression models, etc.

Supplemental Instructor - John Jay College, CUNY

01/2020 - 05/2020

ECO 105: Understanding U.S. Economic Data

- Worked alongside Professor assisting 25 students with classwork, weekly
- Developed lesson plans and learning strategies to accelerate student learning
 - Using data from BLS, BEA, ACS, BJS, Census, World Bank, and Excel
- Facilitated weekly supplemental lessons hosting 3-5 students per session for 11 sessions
- Effectively transitioned from in-person to online assistance due to Covid-19 restrictions
- Improved students overall scores by 12%
- 1 of the 9 students selected to take part in this pilot program inspired by UMKC:
<https://rb.gy/hyvj3q>

Little Curious Explorers - Valley Stream, NY

03/2017 - 03/2020

Manager (03/2018 - 03/2020)

- Oversaw and directed 3-4 employees in their tasks, daily
- Curated company policies and resolved issues to optimize performance
- First Employee of the Month winner, recipient of cash bonus

Event Coordinator (03/2017 - 03/2018)

- Assisted customers and trained 5 new employees
- Created and facilitated company's first Holiday Toy Drive for the Ronald McDonald House Fdn.

Other Experience

Theater Department Fund TDF — Scholarship Recipient

2016 - Present

Museum of the Moving Images — Intern

Oct 2016 - Jan 2017

- Member of museum's first Teen Council.
- Collaboratively curated a social and educational event for 300 intercity teens in the NYC Area.

Amnesty International — Satellite Member

2014 - Present

Education

General Assembly - Data Science Immersive Student

2022

- 10 week, 500 hour intensive Bootcamp covering Python, SQL, Statistical Programming, Machine Learning, and Neural Networks.

John Jay College of Criminal Justice CUNY

09/2017 - 05/2021

- Bachelor of Science, Major in Economics, Double Minor in Fraud Examination and Sociology
- Major GPA 3.95 | Cumulative GPA 3.85
- Magna Cum Laude; Dean's List; Omicron Delta Epsilon Intl Honor Society in Economics