Mark Torres

✉ mark.torres@yale.edu   📞 239-300-3122   in linkedin.com/in/mark-p-torres/   github.com/mark-torres10

Education

Yale University, New Haven, CT   *Expected Graduation: May 2020*
Bachelor of Science, Statistics & Data Science
GPA: 3.59 / 4
Relevant Coursework:
- Data Mining and Machine Learning, Unsupervised Learning and Big Data, Computational Tools for Data Science, Biomedical Data Mining and Machine Learning, Multivariate Statistics, Probability Theory, Theory of Statistics, Discrete Mathematics

Research Publications:
- Iwamoto, S.K., Alexander, M., **Torres, M.**, Irwin, M. R., Christakis, N. A., & Nishi, A. (2020). Mindfulness Meditation Activates Altruism. Scientific Reports, 10(1), 1–7. https://doi.org/10.1038/s41598-020-62652-1

Experience

Crockett Lab, Yale University | Data Science Researcher December 2019 – Present
- Used data augmentation methods to improve the classification performance of a recurrent neural net (RNN) by 26.71%, to a new accuracy of 92.55% on test data.
- Developing a pipeline to automatically stream tweets using the Twitter API, store in AWS, classify with a pre-trained deep learning model, and automatically send messages to users in order to gather further data.
- Experimenting with data augmentation methods (e.g., word embeddings, translation with Google Translate API, self-labeling) to artificially increase training labels.
- Combining deep learning results with focus groups to learn about how people perceive moral outrage.

Groupwise (Yale University startup) | Software Engineer March 2020 – Present
- Developing a pipeline to automate data collection, storage, and analysis.
- Performing statistical analyses, in R, in order to study people's ability to detect mistakes in repetitive tasks.
- Automating analysis and presenting information to stakeholders via Shiny dashboards (in R).

The D. E. Shaw Group | Data Science Intern, Generalist Program May 2019 – August 2019
- Performed statistical analyses, designed algorithms, and implemented ML models in R, Python.
- Implemented data science pipeline. Wrote scripts to automate data analysis for future users by pulling data using SQL queries and wrangling in R and Python.
- Developed a program to parse and standardize existing free-text data using NLP. Successfully parsed 40% of previously unstandardized text data.

Yale School of Management | NLP Undergraduate Researcher September 2019 – January 2020
- Trained neural networks to classify the sentiment of tweets, in order to predict how the opinions of certain people change over the course of their interactions with others online, in particular with automated accounts/bots.
- Scraped tweets using Python to interface with Twitter API. Performed data wrangling and text cleaning.

Human Nature Lab, Yale University| Data Science Researcher September 2018 – January 2020
- Designed and executed experiments to show that meditation resulted in an increase in giving, leading to a research publication.

Skills

Computer Languages: Python (proficient), R (proficient), SQL (intermediate), Java, C++, HTML, CSS, JavaScript
Frameworks: Tensorflow, Keras, PyTorch
Other: Spark, AWS, Git, Tableau, Stata

Achievements

DataFest 2019 (hosted by DataCamp and RStudio) March 2019
- Won "Most Innovative Approach" award by presenting a novel approach to evaluating a rugby team's performance using geospatial data.

Codecademy (online class platform) March 2020 - Present
- Completed 200 hours of computer science coursework (e.g., programming, recursion, data structures, algorithms), implemented in Python.