

# NAVRAJ NARULA

**Email:** nnn2112@columbia.edu

**Website:** <http://navierula.github.io>

**Twitter/Github:** NavieRula

## DATA SCIENCE PROJECTS

### Detecting Fake News Headlines

Created binary questionnaire with open-ended forum asking participants to differentiate between real and fake news headlines. Visualized and reported on collected results against KNN classifier built using R.

### Generating Poetry with Markov Chains

Devised baseline and advanced algorithm to generate coherent text based off haikus in Python, conducted 74-participant survey that demonstrated an error rate of < 30% on selected outputs.

### Analyzing Language in Eating Disorders

Utilized LIWC to obtain pertinent categories in reddit posts related to anorexia and obesity, plotted correlations using R. Built classifiers using sci-kit learn, returning accuracies of 90% for anorexia-based texts and baseline for obesity-based texts.

### Localizing Travel via a Web Application

Built website allowing users to create profiles and list ideal travel destinations using Python, SQL, HTML, and CSS. Displayed data based on user-generated content.

## SKILLS

### Programming

Python - Java - C - HTML - CSS - JavaScript - R - SQL - Git - Flask - D3.js - Shell - pandas - nltk - OpenRefine - Elasticsearch - Praat

### Research/Concepts

Natural Language Processing - Speech and Audio Evaluation - User Experience Survey Design - News and Social Sharing - Technical Documentation

### Languages

English - Thai - Punjabi

### Leadership

Community Chair - Women in Computer Science at Columbia University

## EXPERIENCE

### Software Engineering Intern

June 2018 - Present

*Axios - New York, NY*

Building data pipelines to gather social engagement data using Python and AWS Glue as a storage framework for extract, transform, and load tasks. Creating and deploying machine learning models with AWS SageMaker. Working closely with the Director of Data Engineering and reporting directly to the CTO.

### Artificial Intelligence Researcher

Sept 2017 - Dec 2017

*Creative Machines Lab - New York, NY*

Designed and implemented a persistence algorithm and an autoregressive model using time-series forecasting to predict future movements of a soft, artificial muscle. Visualized muscle movements pertaining to each time second and load force for continuous datasets.

### Software Engineering Intern

June 2017 - Aug 2017

*Quid - San Francisco, CA*

Built and deployed website to display info pertaining to Quid's backup system using Docker, Flask, SQLAlchemy, and Python. Examined and unpacked contents of API to create mapping and define settings for indices on Elasticsearch. Wrote test cases to automate checks for an infrastructure engineering task. Presented my work and received recognition for delivering outputs of my code to a client.

### Teaching Assistant

Sept 2016 - May 2017

*Columbia Engineering - New York, NY*

Helped students understand control flow, functional design, and complex libraries using Python. Selected as 3 of 10 course assistants to lead lab lectures.

### Digital Products Intern

May 2016 - July 2016

*MIT Press - Cambridge, MA*

Implemented functions in Python script to automate ePub creation, validated parsing errors across XML/HTML documents, bookmarked PDF textbook files, worked with Oxygen XML Editor and Sigil-Ebook platform to run queries and automate test cases.

### Data Storytelling Researcher

Jan 2016 - May 2016

*BU Hariri Institute for Computing - Boston, MA*

Designed taxonomy and implemented a relational database in MySQL Workbench to store info about criminal case reversals, worked with faculty member to deliver report on functionality of search and how to perform robust queries, received funding to attend NICAR conference in Spring 2016.

## EDUCATION

### Columbia University, New York, NY

May 2018

*Master of Science, Computer Science and Journalism*

Thesis: Detecting Human Emotion in Speech and Audio Inputs

### Boston University, Boston, MA

May 2016

*Bachelor of Science, Computer Science and English Education*