

# Ruizhen Mai

rm4610@nyu.edu | 929.855.2208 | 626 1st Ave. Apt. e30h, New York, NY 10016

## EDUCATION

### New York University

*Bachelor of Arts, Major in Computer Science and Mathematics, Minor in Business Studies*

**New York, NY**

*(Expected) May 2021*

- Major GPA: 3.79 / 4.0
- Dean's List all semesters
- Relevant Coursework: Machine Learning, Basic Algorithms, Parallel Computing, Operating Systems, Network systems, Computer Security, Big Data, Linear Algebra, Probability Theory, Statistics, Real Analysis, Stochastic Process, Time Series

## TECHNICAL SKILLS

Web and Big Data: React, Node.js, MongoDB, MySQL, Bootstrap, Html, Spark, Hadoop, Kafka

Machine Learning and Data Analytics: PyTorch, TensorFlow, CUDA, Pandas, Matplotlib, Plotly, Selenium

Programming Languages: Python, C++/C, JavaScript, Golang, R, Java, SQL

## EXPERIENCE

### New York University

*Research Assistant w/ Prof. Xinyu Cao (Python)*

**New York, NY**

*Nov 2019 – Present*

- Developed experiments to try identifying optimal predictors to auction prices of artwork, jewelries and furniture
- Built and integrated machine learning models for image, text, categorical and continuous data

*Research Assistant w/ Prof. Joshua Loftus (Python, R)*

*May 2019 – Present*

- Created NeuroMICE, an improvement of MICE algorithm (multiple imputations by chained equations) that can dramatically reduce time needed for specification in the algorithm and the barrier to use
- Developed and streamlined the experiment to test NeuroMice to accomplished preliminary results that demonstrated comparable 95% confidence interval coverage
- Assessed the strengths and weakness between Joint Modeling (e.g. EM) and Fully Conditional Modeling (e.g. MICE)

*Research Assistant w/ Prof. Tommaso Bondi (Python)*

*Apr 2019 – Oct 2019*

- Collected and organized 400K book data from Goodreads and Bookmarks using **Selenium** and **Pandas**
- Researched the differences between the reading preferences of Goodreads new and experienced users, and performed t-test to discover that readers' leniency on ratings and the range of preferences decrease as they become experienced
- Built neural networks to classify sentiments of consumer reviews with 83% accuracy

*Project Assistant w/ Prof. Xiao Liu (Python, R)*

*May 2019 – Aug 2019*

- Investigated and confirmed the results of "Retention Futility" (Ascarza, E. 2018) – retention campaign targeting high-risk customers is ineffective, while aiming at "lift" is more appropriate – with causal forest

### uStart.today

*Software Engineer Intern (Golang, JavaScript)*

**Brooklyn, NY**

*Mar 2018 – May 2018*

- Programmed an alpha-version chat API using **WebSocket** and handled the message emission in the front end

## PROJECTS

**Real-Time Stock Price Forecasting Application** (React, Node.js, Python, R)

*Jan 2020 – Present*

Developed an application that displays and updates a stock price and its forecasts every minute

- Developed a **Python-R** program to forecast the stock price, its volatility, and its option prices using ARIMA-GARCH model and Black-Scholes formula
- Constructed a REST API using **Node.js** to retrieve stock data from worldtradingdata.com and send them to front end
- Designed a responsive website using **HTML5**, **Bootstrap** and **React** to display the stock price

**CycleGAN w/ Color-Shape Regularization** (Pytorch)

*Apr 2019 – May 2019*

Modified CycleGAN model with regularization that punishes generated image's incorrect shape

- Brainstormed the idea by analyzing the flaws in the original model
- Programmed the CycleGAN model, and tested the regularization with 1000 images from horse2zerbra dataset

**Poverty Rate Analysis** (R)

*Apr 2019 – May 2019*

Searched for correlated and useful linear predictors to poverty rates

- Collected poverty rate data, and possible predictors data such as unemployment rate, birth rate from World Bank database
- Built regressions of poverty rate on these possible predictors, and analyzed their performances by r-squared, p-values of each predictor's t-test and diagnostic plotting of the residuals from the regression
- Discovered that percentage of population having access to electricity as a strongly correlated predictor to poverty rate with 0.8 r-squared in a single-predictor model and over 99% to include it in a six-predictor model