Elsa(Jirou) Xu

31 River Court, Apt 1109
Jersey City, NJ, 07310
646-705-4996
jx654@nyu.edu

Seeking a Full-time Data Scientist Job Opportunity

Education

Sept. 2015 - present **New York University - Master of Science in Data Science**.
Core courses: Statistical and Mathematical Methods, Big Data, Machine Learning, Natural Language Processing, Deep Learning, Causal Inference
Graduation May 2017

GPA **3.62/4.0**.

Project Experience

February– March 2017 **MNIST handwritten Digit Recognition**, *New York University*.
Implemented two different models Variational Autoencoders (VAE) and ConvNet to do the handwritten Digit recognition on MNIST datasets. Improved final result to 99% by using data augmentation and pseudo labels technique and adding layers to ConvNet. *Python environment*

October– December 2016 **Deep Learning for Story Comprehension Tests**, *New York University*.
Applied GRU layers and dynamic memory network model on Who-did-What dataset, a large scale person-centered cloze data set with multiple choice question. Improved the accuracy of multiple choice answer by using n-gram word representation and changing the memory update mechanism of dynamic memory network model. *Python environment*

April– June 2016 **Sentimental Analysis on Movie Reviews of IMDB** , *New York University*.
Did sentimental analysis on a multi-label movie review corpus. Different from traditional text classification, texts are labeled by several sentiment labels. Chose SVM as the baseline model. Other methods included Naive Bayes, logistic regression, k-Nearest Neighbors, and Random Forest. Used a mixture of unigrams and N-gram. *R environment*

April– June 2016 **Automated Question Answering System of Yahoo Answers** , *New York University*.
Aimed at searching similar question and providing the corresponding answer on the Internet-based knowledge exchange website Yahoo!Answers when a new question is asked. Improved the accuracy of determining categories of questions by 6% using Ensemble of multinomial Naive Bayes model. Found the final answer using Okapi BM25 model. *Python environment*

March– May 2016 **NYC Taxi Data Merge and Join** , *New York University*.
Used map-reduce jobs to deal with NYC taxi data, focusing on joining the trips and fares data together, extracting important information from data of different taxi and driver. Mainly used the knowledge of key-value pair to do the map-reduce jobs. Set different features of NYC taxi data as key and defining their corresponding values. Ran code on both AWS and Hadoop to get a general result. *MySQL environment*

Professional Experience

July – August 2016 **Assistant Data Analyst**, *Overseas Chinese Fund*, Hangzhou, China.
Wrote SQL queries to get age, investment amount and geographic distribution for investors throughout the whole year. Used collected data to draw a line chart to demonstrate the results of the survey to help the company in:

- Finding potential investors of all ages and further expanding the business regions.
- Determining the peak of investment every day to make proper arrangement. Increasing the capacity of data traffic by 30% during the peak hour.
- Finding the most popular range of investment amount and adjust the investment amount choices. Eliminating 3 redundant options.

Computer skills

Intermediate LaTeX, Microsoft Windows, HTML

Advanced MySQL, MATLAB, Python, C/C++,R