JIALI ZHOU

45 RIVER DRIVE SOUTH, JERSEY CITY NJ,07310

+1(551)208-4678 ⋄ jz2312@nyu.edu

OBJECTIVE & INTRODUCTION

Seek **summer internship** to utilize my rich experience and background in **data science** and **machine learning**. Passion for big data practices, new technology as well as sharing knowledge and experience and creating useful data insight and application.

EDUCATION

New York University, Courant Institute of Mathematical Sciences, New York, NY   May. 2017
Master of Science in Data Science, focus on statistics and machine learning
· **Related Courses:** Machine Learning, Natural Language Processing, Statistical and Mathematical method, Artificial Intelligence

Tsinghua University, Beijing, China   Jul. 2015
Bachelor of Science in Electrical and Computer Engineering
· **Overall GPA:** 3.80/4.00
· **Related Courses:** Computer Vision, Data Structure, Real Analysis, Introduction to Artificial Intelligence, Numerical Analysis, Graphic Models
· Academic Individual Scholarship at Tsinghua University in both 2013 and 2014

SKILLS & STRENGTHS

Statistical capabilities	Linear/Logistic Regression, Decision Tree, Adaboost, Bayes, SVM, KNN, etc
Computer Languages	Python, C/C++, R, MATLAB, Ruby, Javascript, php, etc.
Database	Mysql, Data Mining and In-Database Analytics with SQL, MapReduce, etc
Languages	Mandarin (mother tongue), English (fluent)

RELEVANT EXPERIENCE

Baby name popularity analysis provided by the Social Security Administration   Capital One, Mar.2016
· Using logistic regression to predict the name that that may have had the largest increase or decrease in popularity.
· Utilized data visualization to present the high dimensionality of features of state, name and year using python and xml.

Air Flight Delay Prediction   New York University, Oct.2015 to Dec.2015
· Extracted five main reasons for the air flight delay and then gather the related data such as flight and weather data, etc.
· Selected informative features by lasso regression and Principle Component Analysis and implement SGD SVM, Decision Tree, Logistic Regression and Random Forest on our data set to train the model and got 91.1% accuracy.
· Used data visualization to present the high dimensionality of the correlation of weather, airport control and tailnum features with the delay variable.

Supporting Search-As-You-Type Using SQL in databases   SAP Company, Dec.2015 to June 2015
· Presented solutions for both single-keyword queries and multi-keyword queries, and improved search using SQL by allowing mismatches between query keywords and answers.
· Presented techniques to answer first-N queries and discussed how to support updates efficiently.

Analysis of motifs in the human gene   Tsinghua Univeristy, Summer 2014
· Found the most suitable database NATsDB from different papers and websites, extracted and processed the data from the selected database, and applied the DME algorithm in pattern recognition methods to conduct experiments.

Data Mining for Likelihood associated with Acute Hospital Readmission   Purdue University, Summer 2013
· Used statistical models (logistic regression, decision tree) and improved mixed models to predict readmission rates, which improved accuracy by 5%.
· Submitted paper: Developing Conditional Logistic Regression Models for Predicting 30-Day Hospital Readmission.

HONORS & AWARDS

· **Meritorious Winner Award in the 2013 Interdisciplinary Contest in Modeling in 2013**,
· **Second Prize in the National Undergraduates Physics Competition in 2012**,