# Guifan Li

*178W 171th St. Apt 33, New York, NY 10032 • 917-847-0450 • gl2439@columbia.edu*

## EDUCATION

**M.S**, ***Columbia University –*** Computer Science, GPA 3.55/4.0 *May 2014*
**B.S**, ***University of Melbourne*** – Mechatronics Engineering/Computer Science, GPA 3.6/4.0 *Feb 2012*
***Student Ambassador, University of Melbourne***

## TECHNICAL SKILLS

Java, C++/C#/C, Python, SQL, Matlab, R, PHP/CSS/HTML, AWS, Linux, MS Office, VBA, Bloomberg API.

## PROFESSIONAL EXPERIENCE

**Trust Company of the West (TCW)** **New York, NY**
*Financial Engineer Intern – Emerging Market Group* *Jul 2014 – Sep 2014*

- Assisted a $6 billion asset group to design and build a database in MS SQL to handle up to 10 million rows of data in several tables and built models to house equity prices, financial estimate, ratios and growth rates in time series.
- Optimized the database with Indexes, Views and Store Procedures to decrease processing time from 30 to 4 hours.
- Established a system using Java through Bloomberg API to allow the database automatically update historical data of stocks in two portfolios weekly and output the processed data to Excel in VBA.
- Enabled the team to analysis data according to countries, sectors, sub-sectors and combination of them in a historical perspective to explore implied correlation between stock price and those financial parameters.
- The system is heavily used on a daily basis and my intern performance evaluation was "Exceptional" (top ranked).

**Crowdnetic** **New York, NY**
*Data Engineer Intern, R&D Team* *Mar 2014 – May 2014*

- Designed and built a system to classify customers' credit into 7 grades with each class containing 5 sub-grades based on 7 features in their profile with C++, improving accuracy of old classifier from 55% to 91% by creating a new algorithm after exploring the features of the data sets.
- Applied the classifier on data-driven products such as a recommendation system, credit grade, loan and interest rate predication models for P2P market and applied big data/advanced analysis to identify and exploit insights.

## ACADEMIC EXPERIENCE

**An innovative Multilayer Neutral Network algorithm to classify big data (Python)** **Columbia University**
*Research Assistant* *Jan 2014 – May 2014*

- Developed and implemented an innovative algorithm, which combined Stochastic Gradient Decent (SGD) and Ada-boosting algorithms to classify big data and improved the efficiency and accuracy of traditional algorithms.
- Decreased computational time of the traditional algorithm in Artificial Neutral Network by 1/3 and increased accuracy from 88.7% to 97.64% for MNIST - a handwriting detection dataset (world record is 98.47% but slower).
- Incorporated data mining tools, such as Vowpal Wabbit, Ski-learn and Orange and made comparison among them.

**Basketball data analysis center website (Java/PHP/CSS/Oracle)** **Columbia University**
*Team Leader* *Feb 2014 – May 2014*

- Designed the detailed schema and structure of complicated database system, including E/R diagram and relationships.
- Built a website using PHP/CSS and Oracle to show performance and analysis of players, teams in past seasons according to the data retrieved from NBA website in real time using a web crawler implemented in Java.

**Online dating application: Dategram (Java/JSP/Servlet/MySQL)** **Columbia University**
*Team Leader* *Feb 2014 – May 2014*

- Built an IOS app, utilized Amazon AWS as back-end, and designed front to back-end interaction prototype.
- Designed E/R diagram, implemented MySQL database and developed functions in Java such as signing up, adding friend, and dynamic self-updating information, ranking people, and filtering out recently unmatched people.
- Implemented a matching algorithm to connect different people via analyzing their Big Five personality and common interests posted via social media.

**Machine Learning and Data Mining course projects (Matlab/R)** **Columbia University**
*Machine Learning Projects* *Sep 2013 – May 2014*

- Amended the KNN algorithm to classify dataset on 9 features with an averaged accuracy of 83% vs. 70% benchmark.
- Built a Gradient Decent algorithm to classify 2 classes dataset with an accuracy of 96% vs. 80% benchmark.
- Applied a Naïve Bayes algorithm to build up a junk mail binary classifier based on their texts with an accuracy of 91%.

## LEADERSHIP EXPERIENCE & AWARDS

**Teaching Assistant**, Advanced Machine Learning, Columbia University *Jan 2014 – May 2014*
**7th place at Melbourne site**, ACM-International College Programming Contest (ICPC) *Sep 2008*