**Jiaxin Li**

536 West 113th Street, Apt 4, New York, NY 10025

[jl5681@columbia.edu](mailto:jl5681@columbia.edu) | (332) 207-6051

<https://www.linkedin.com/in/jiaxin-li-3262221a3/>

**EDUCATION**

**Columbia University,** *M.S. in Applied Analytics*, *New York, NY* June 2021

* **Coursework:** Natural Language Processing, Data Mining, Data Science Consulting, SQL & Relational Database,

Machine Learning, Monte-Carlo Simulation, Principal Component Analysis

* **Computer:** Python, Java, R, SQL, Tableau

**Central University of Finance and Economics,** *Bachelor in Electronic Commerce*, *Beijing, China*July 2019

* **Relevant Courses:** Introduction to Probability and Statistics, Advanced Programming and Algorithms, Econometrics, Operation Research and Methods (Major GPA: 3.7/4.0)
* **Honors:** Excellent Student Scholarships (Top 5%: sophomore & senior), Excellent Scholarships for Internship (three consecutive years), College Student Entrepreneurship Project Excellence Award of CUFE (2017)

**RELEVANT EXPERIENCE**

**Hirebeat Co., Ltd.** New York, NY

*Data Scientist* Aug 2020 – Nov 2020

* Predicted high-frequency keywords from LinkedIn/Indeed job descriptions by implementing machine learning models (SVM, logistic regression, random forest) and neural networks (MLP and CNN) with 96.7% fit.
* Extracted keywords from job descriptions that were scraped from LinkedIn/Indeed using various Natural Language Processing (NLP) methods.
* Created AWS Code Commit repos for the projects for later execution and pushing the codes.

**China Securities Co., Ltd.**Beijing, China

*Data Analyst* Oct 2018 – Feb 2019

* Created a Python script to connect to WIND financial terminal API and retrieved historical stock data of brokerage and insurance firms. Generated automated reports using Excel macros, reducing overhead time from two hours to minutes.
* Calculated investment potential financial ratios (e.g., P/E, P/B, and EPS) and compared and visualized competitor performances using Tableau. Assisted senior analysts in making investment recommendations (Buy/Sell/Hold).

**AXA China Region Insurance Co., Ltd.** Hong Kong

*Data Analyst, Market Research Department*July 2017

* Automated the daily patent data query from OTD’s official API. Cleaned and inserted the data into firm’s SQL database.
* Built a custom valuation system for different patent categories and implemented gradient boosting, neural network, decision tree, RNN and LSTM methods to predict patent value with 95% fit.
* Implemented generative adversarial networks to generate statistically similar synthetic data for sample size increase.

**PROJECTS**

**Kaggle Project for COVID-19 Diagnosis from Pneumonia Cases’ Chest X-Rays**

* Used the Keras ImageDataGenerator to resize the image to 224x224x3 and applied various image augmentations (zooms, horizontal flips, width shifts, height shifts, brightness changes, rotations) to detect virus particles.
* Applied transfer learning and fine tuned large pre-trained CNNs such as DenseNet and concluded that ~85% accuracy was the best a single CNN could perform.
* Combined transfer learning with Ensembling and finetuned 10 large pre-trained networks like VGG16, DenseNet121, DenseNet169 and ResNet50 and recorded prediction probabilities from each model. I averaged the probabilities for each sample to get the "group-voted-best answer", which improved the generalization of my final predictions.

**Data Science Consulting project for Retail Clients**

* Built a monthly report for retail clients based on their customers’ data, customers’ use of retail websites and purchases made. Analyzed popular products and targeted valuable customers, and communicated with product managers based on the results for decision making of next season stock-up.
* Created a user-interactive website using Shiny for target selection and dynamic graphic illustration.