

# BOWEN ZHANG

808 Columbus Ave, New York NY 10025 • (917) 547 3523 • [bz2461@columbia.edu](mailto:bz2461@columbia.edu) • LinkedIn

## EDUCATION

### Columbia University

*M.S in Business Analytics*

New York, NY  
Expected Dec 2023

- **GPA:** 3.92/4.0
- **Relevant Coursework:** Artificial Intelligence, Data Analytics, Business Analytics, Optimization

### University of British Columbia

*B.Comm in Business and Computer Science Honors*

Vancouver, BC  
May 2022

- **GPA:** 3.85/4.0, Dean's List
- **Relevant Coursework:** Machine Learning, Statistics, Advanced Database, Algorithm Design

## WORK & RESEARCH EXPERIENCE

### Noumena

*Data Science Intern*

New York, NY  
Jan 2023 – May 2023

- Productionized the **regression** model to find intrinsic values for self-employed workers based on the survey information provided; leveraged the model to estimate clients' potential and achieved **RMSE** of 3.23
- Reported the most important survey questions by drawing out **SHAP** force plot and summary plot; cooperated with business and tech teams to integrate the model into online services

### China International Capital Corporation (CICC)

*Data Science Intern*

Beijing, CN  
Jul 2021 – Oct 2021

- Developed ML models to assess the likelihood of loan default based on financial metrics from the WIND database and customer demographic data; achieved **roc\_auc** score of 78% and contributed to \$0.2M potential annual savings
- Implemented and compared algorithms namely **Random Forest**, **XGBoost**, and **LGBM** to make predictions; built pipeline for data preprocessing, conducted hyperparameter tuning using **RandomizedCV** with 5-fold **cross-validation**

### SAP

*Business Analyst Intern*

Vancouver, BC  
Jan 2020 – Sep 2020

- Designed an interactive tool to display financial metrics of target firms, including DIO, DPO, and COGS, and benchmarked with peers to facilitate internal sales teams to target potential customers, received 300+ downloads
- Wrote **SQL** queries to retrieve and cleanse financial data from S&P; implemented **time series analysis** to showcase firms' financial performance utilizing **R** and Business Intelligence software **SAP Analytics Cloud**

## TECHNICAL PROJECTS

### Deep Learning: Loan default prediction

Apr 2023 – May 2023

- Built a credit scoring model by predicting loan default rates with 3-layered **feedforward neural networks** for LendingClub in **Keras**; optimized the model with **stochastic gradient descent** and achieved accuracy of 90%
- Implemented **EarlyStopping** and **dropout** to avoid overfitting; added **batch normalization** for data scaling; tested the model across 500 epochs with different batch sizes to improve accuracy

### Natural Language Processing: Yelp Review

Apr 2023 – May 2023

- Predicted Yelp business reviews based on user comments by comparing models such as **DistilBERT** and **MultinomialNB** to improve the baseline accuracy by 11%; conducted **LDA** and **sentiment analysis**
- Applied various preprocessing methods to machine learning models namely tf-idf, tokenization, and PorterStemmer

### Big Data: Movie Recommendation

May 2020 – Jun 2020

- Built a **collaborative-filtering** movie recommender system with **PySpark ALS**, leveraging the computing power of **AWS EC2**; predicted movie ratings with RMSE of 1.36

## SKILLS

- **Quantitative Skills:** proficient in Python, R, SQL, Java, Spark, Machine Learning & Neural Networks, Excel
- **Language:** English, Mandarin