

# 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

- **Relevant Coursework:** Business Analytics, Probability and Simulation, Data Analytics

### 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

### China International Capital Corporation (CICC)

*Data Science Intern*

Beijing, CN

Jul 2021 – Oct 2021

- Developed ML model to assess the likelihood of loan default based on financial metrics from the WIND database and internal 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

### University of British Columbia

*Research Assistant, Supervisor Professor Jackie Stewart*

Vancouver, BC

May 2021 – Sep 2021

- Collected survey data to analyze how students' emotions impact their learnings; analyzed data leveraging Python and R and presented findings in the Canadian Chemistry Conference and Exhibition 2021
- Conducted exploratory factor analysis to extract 5 key features from the survey result, performed Two-Way ANOVA test to find the most important feature affecting grade, and built Structural Equation Model to test the inference

### 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

### Spark Streaming: COVID Analysis

Aug 2022 – Sep 2022

- Processed real-time COVID data by connecting to Twitter API with tweepy, conducted sentiment analysis leveraging NLTK sentiment analyzer; concluded that people are positive about the current COVID situation

### Deep Learning: Loan default prediction

Jun 2022 – Jul 2022

- Built a credit scoring model by predicting loan default rates with 3-layered deep neural networks for UCI Credit Card, utilizing Tensorflow; optimized the model with stochastic gradient descent and achieved accuracy of 90%

### Natural Language Processing: Yelp Review

Sep 2021 – Oct 2021

- Predicted Yelp business reviews based on user comments leveraging MultinomialNB classifier in Python and achieved 92% accuracy; created pipeline that converted strings into feature matrix with tf-idf

### 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
- **Libraries:** Numpy, Pandas, Matplotlib, Seaborn, Scikit-Learn, PySpark, MLlib, NLTK
- **Language:** English, Mandarin