# Resume

## Contents

Chenshuo Pan

Address: | Tel: 608-628-6056| Email: cp3384@cumc.columbia.edu

## EDUCATION

**Columbia University**    New York, NY

Master of Science in Biostatistics

- Coursework: Biostatistics, Data Science , Probability , Epidemiology

**University of Wisconsin–Madison Madison**,    WI

Bachelor of Science in Statistics

- Coursework: Data Model, Deep Learning Probability and Math Statistics, Stochastic process, Appl Categorical Data Anal, Data Science with R, Machine Learning

- Honours: Dean's List(2021/2022/2023)

## PROFESSIONAL EXPERIENCE

**Baidu Online Network Technology (Beijing) Co., Ltd**    machine learning intern Jun 2023 – Aug 2023

- Create internal learning materials related to the basics of large models and explain them to sales staff

- Read literature and assist in the production of internal learning materials including llama, chatglm, falcon and other open source models and multi-modal models

- Assist in completing llama2 related technical analysis, and test the deployment, fine-tuning and other practices of llama2 based on Baidu's Qianfan platform

- Participated in the enterprise privatization model deployment project, using Baidu's Wenxin model as a basis to customize large industry models for many enterprises

**CSC Financial Co., Ltd**    Investment Banking Intern Jun 2022 – Aug 2022

- Participated in the preparation of relevant documents for multiple development bidding of the department, assisted in writing relevant bidding documents, and completed the production of bidding PPT

- Collected and extracted financial reports and statements of multiple Chinese listed companies in the past three years using Wind Financial Terminal and other data platforms

- Analyzed the post-listing operations of state-owned and private enterprises and their pattern, and drafted reports

- Assisted in the completion of the feedback and reply of a project in the group to the questions inquired by the CSRC

**Baidu Online Network Technology (Beijing) Co., Ltd**      Internet Finance Intern Mar 2022 – May 2022

- Analyzed the current intelligent push, credit evaluation and some other functions and mastered the operation mode and infrastructure logic of Baidu's big data products

- Analyzed Baidu Cloud and cloud finance, learned the relationship and practical use case of cloud and AI

- Analyzed the data push algorithm of Little Red Book (a social media and ecommerce platform), and analyzed the big data ecology and components used, and gave industry development prediction

**China Telecom Hongxin Information Technology**      Developer Intern Jun 2020 – Sep 2020

- Responsible for front-end and back-end interface development, web page and program design

- Participated in the function development of WeChat public account, and launched a small program smart community beta version

- Wrote a intelligent community program interface using JavaScript, ran and tested the program afterwards

- Analyzed project documents and plans, and responsible for the development of back-end functional interfaces based on the spring boot framework


## ACADEMIC PROJECTS AND PUBLICATIONS

**Research on the technical architecture and key technologies of data element value**   Jul 2022 – Dec2022

- Researched on metadata related literature, analyzed the characteristics, and compared the advantages and disadvantages of metadata management systems used by Apache, LinkedIn, and other companies

- Assisted in building a metadata-based knowledge graph prototype system for team project

- Compared and adjusted the existing algorithms including linear regression, logistic regression and LGB to obtain higher level of accuracy based on the actual business scenario data, with the goal of detecting outliers/sequences

**Analyzing the Salary of NBA's Player Data Project**   Sep 2021 – Dec 2021

- Collected over 450 pieces of NBA player data for different position, and visualized the data using ggplot2, sunburstR

- Collected salary data from Kaggle, analyzed and visualized player salary distribution using Python Information Classification: General

- Analyzed various factors including age, draft order, performance and identified the main influencing factors of salary, and offered suggested salary to both team manager and players accordingly

- Built a database and trained data for the use of machine learning algorithms including linear regression, classification, and regression tree to predict salary of NBA's players, using the optimal algorithm to build a well-designed compensation system

**Research on Metadata Management of Data Lake published on SPIE - The International Society for Optical Engineering(Submitted for Index in EI Compendex & Scopus)**

## SKILLS/INTERESTS

Computer Skills: Python, R, Java, C(base), JavaScript, Excel, SQL

Languages: Mandarin (native), English (fluent)

Interests: Basketball, Calligraphy