

Jingjie (Jennie) Zhou

🏠: New Jersey 📞: (917)855-2658 ✉️: zhoujingjie06@gmail.com 🔗: linkedin.com/in/jingjiezhou/

SUMMARY

Diligent Data Scientist with strong statistics, math and finance background delivering data-driven insights to improve business efficiency. Excellent understanding of analytics tools experienced at machine learning models and interpreting large dataset.

WORK EXPERIENCE

Data Scientist

NY, Feb 2021 - Now

QuaEra Insights

- Analyzed and improved the effectiveness of client's email campaigns by building machine learning models.
- Designed algorithm to identify multiple content blocks, scraped all text and image from Email campaigns web pages (HTML).
- Identify and extracted address elements from direct mail using **NLP** technique. Elements are further processed as features in the pipeline for the model.

Data Scientist Intern

New York Life Insurance Company

NY, June 2020 – Aug. 2020

- Ingested large dataset from disparate data source using complex **SQL** queries.
- Defined, developed and evaluated stages of data-driven product ETL optimization in **R** scripts based on functional design.
- Optimized survival-analysis-based Automatic Underwriting Process with underwriting team members by 10%.
- Guided and improved ETL process for clients' lab test result considering both statistical and business aspects.

Deep Learning Research Intern

Fordham University Collaborate with NYU Langone School of Medicine

NY, May 2020 – Aug. 2020

Deep Convolutional Neural Networks for Predicting Head Pose During Brain MRI Acquisition (Published on VSS'20)

- Designed all stages of artificial intelligence implementation on MRI scan process with NYU medical researchers.
- Defined and developed innovated consecutive video image processing strategies in **Python** scripts.
- Developed CNN model for GPU computing. Evaluated the performance by visualization and statistical metrics.
- Successfully predicted head pose motion, reduce repetitive operations, boosted medical equipment efficiency by 60%.

Corporate Finance department Bank Manager

Bank of Ningbo

Beijing, Jan. 2016 – Apr. 2018

- Integrated Corporate Finance Datasets from several data sources.
- Created visually financial report by using advanced **Excel** Pivot tables and **VLOOKUP** functions.
- Extracted, interpreted, analyzed data and prepared reports of potential company clients for loan decision making.

PROJECTS

2020 US Election Tweets Analysis

Aug. 2020 – Oct. 2020

- Designed and developed the pipeline. Utilized **Python** data extraction and manipulation, implemented **NLTK** library to **NLP** data processing and creating various Data Visualization Dashboards to analyze tweets.
- Categorized tweet comments into positive and negative clusters and created WordCloud using Sentiment Analysis.
- Visualized the sentiment changes over time before election day and the differences between tweets contents.
- Constructed N-Grams and **LSTM** language models to generate text for different sentiment.

Movie Recommendation System

Mar. 2020 – May 2020

- Built Movie Recommendation System on Databricks (**Apache Spark**).
- Implemented **ALS-Matrix Factorization** recommendation algorithm, tuning the model parameters by **Spark ML** grid search and cross-validation techniques. Manipulated and Visualized user's movie rating prediction.
- Built movie recommendations webpage by **HTML**, **CSS**, **JavaScript** to provide "you might like" similar movies for users.

EDUCATION

Fordham University (GPA 4.0)

Aug. 2019 -Dec. 2020

Master of Science in Data Science (TA for Python Course)

Illinois Institute of Technology

Aug. 2013 - May 2015

Master of Mathematical Finance

Wuhan University

Sept.2009 - July 2013

Bachelor of Engineering in Financial Engineering

SKILLS AND PROGRAMMING LANGUAGES

Programming Language: Python, MySQL, MS-SQL, R, JavaScript

Python Packages: Numpy, Pandas, Matplotlib, Seaborn, scikit-learn, Keras, TensorFlow, CV2, NLTK, BeautifulSoup

Technique: Deep Learning, Machine Learning, Survival Analysis, A/B Testing.

Tools: Advanced Excel, MongoDB, Spark, AWS, Google Cloud, Hadoop, Tableau, Databricks.