Jingyi Zhou

984-215-8391 • Sunnyvale, CA • jingyz2021@gmail.com • www.linkedin.com/in/jingyi-zhou

SUMMARY

A data-savvy storyteller with excellent skills in Python and SQL; extensive experience in data science and business intelligence; efficient partnership with global colleagues on a multi-functional, multi-disciplinary team.

EDUCATION

University of North Carolina at Chapel Hill 08/2018-05/2020
Master of Science in Information Science (UNC Graduate students do not carry a numerical GPA.)

Hefei University of Technology, China 09/2014-06/2018
Bachelor of Engineering in Information Management and Information Systems (GPA: 4.0/4.3)

SKILLS

- **Programming:** Python (Spark, Django), SQL, R, Shell, Java, HTML, CSS (Bootstrap)
- **Tools:** Git, AWS, Docker, MS Office Suite, Power BI, Tableau, SAP BODS, Google Analytics, Scrum, JIRA

EXPERIENCE

Department of Computer Science, UCI Irvine, CA (Remote)
Research Specialist 08/2020-present

- Hosted the genomic visualization web browser for human genome annotation using **Python Django** on **AWS** EC2
- Improved the website loading speed by 340% and reduced the **SQL** query execution time for genomic data by 69%
- Developed single-cell data analysis pipelines on an **HPC** cluster

CommScope, Inc. Hickory, NC
IT Business Analytics Intern 05/2019-08/2019

- Built **ETL** flows for CRM data using **SAP BODS** for data migration and maintained data warehouse architecture
- Designed and automated data visualization dashboards of sales data using **Power BI** and delivered diverse quarterly visualization for different groups in the sales department through cross-functional collaboration
- Increased the accuracy on sales opportunities prediction by 12% by **regression models**, given highly imbalanced data

Neusoft Corporation Hefei, China
Operation & Maintenance Engineer Intern 10/2017-11/2017

- Operated and troubleshot the Electronic Archives Management System, logged and documented tools and procedures to stage and ship client fileservers
- Conducted infrastructure change requests using **Java** based tools and added visualization modules in monitoring tool
- Guided clients in debugging and achieved 100% client satisfaction

PROJECT

Characterizing Allergy Related Comments on Twitter (Python) 02/2020-05/2020

- Crawled and sampled 12 million tweets in 2019 using pre-defined queries, collected weather data of 16 U.S. cities
- Detected time series patterns of tweets count and revealed the correlation between weather and allergy related tweets
- Discovered 152 topics by using the Mallet implementation of Latent Dirichlet Allocation in **Python** and visualized the results with pyLDAvis, t-SNE and bokeh, used Linguistic Inquiry and Word Count to analysis the content of topics

Face Recognition WeChat Mini Program (Python) 01/2020-02/2020

- Trained a face detection model with SSD framework and achieved an average IOU of 77.2% using WIDER face dataset
- Created a Python-driven RESTful API using **Flask** to deploy the well-trained model on the web
- Developed a WeChat Mini Program allowing users to recognize faces in real-time using the API

Sentiment Analysis for Finance News (CO-OP with Credit Suisse) 01/2019-04/2019

- Gathered 400k+ news related to certain stocks from various financial web sources using beautiful soup in **Python**
- Created lexicon-based **sentiment analysis** using Loughran McDonald sentiment word list and NTUSD-Fin dictionary
- Evaluated the performance on the initially selected stocks and achieved 62% accuracy
- Integrated the model and built a real-time sentiment analysis pipeline for live stock news