Wanyi Su

+1 778-319-9915 • wysujhu@gmail.com • Vancouver, BC, Canada

More on Personal Website: https://wanyisu.netlify.app/ • GitHub: https://github.com/suansuan0915

SKILLS

Languages: Python, SQL (MySQL, PostgreSQL), NoSQL (Cassandra), JavaScript (Node.js), Scala, Java, MATLAB, R, bash scripting.
BI and Other Software: Tableau, SVN, Git, Docker, Kubernetes, Control-M, Jupyter Notebook, IntelliJ, MS Office.
Frameworks and Libraries: Spark Scala, PySpark, Kafka, Hadoop; pandas, NumPy, Sklearn, Matplotlib, Seaborn, PyTorch.
Cloud Platform: AWS (EC2, S3, Redshift), Azure.

EDUCATION

Simon Fraser University | Burnaby, BC, Canada   Sep 2021 – Jan 2023
M.S. in Computer Science - Big Data (GPA: 4.0/4.33)

Johns Hopkins University | Baltimore, MD, United States   Jul 2016 – Aug 2017
M.S. in Finance

Southwestern University of Finance and Economics | Chengdu, China   Sep 2012 – Jun 2016
Major in B.A. in Marketing; Minor in B.E. in Finance

WORK EXPERIENCE

Insurance Corporation of British Columbia (ICBC) | North Vancouver, BC, Canada   May 2022 – Sep 2022
Data Engineer (Internship)

- Developed and maintained an ETL data pipeline in **Spark Scala** with 100% accuracy to facilitate data migration.
- Implemented data validation using **SQL** queries to extract large scale data from different databases (**Oracle**, **Drill**, **Hive**).
- Applied software engineering best practice during the pipeline building process, including unit testing, **shell scripting**, CI/CD automation using **Jenkins**, version control using **SVN**, etc.).
- Built **Tableau** dashboards to provide varying levels of visualization for stakeholders and automate data quality monitoring.
- Collaborated with internal and external stakeholders in **Agile** environment (**Scrum** meetings, **JIRA**, **Kanban** board).

SFU Computing Science School | Burnaby, BC, Canada   Sep 2022 to Present
Graduate Teaching Assistant (CMPT 726 - Machine Learning, CMPT 105W - Social Issues and Communication Strategies in CS)

Morgan Stanley | Baltimore, MD, United States   Jan 2018 – May 2018
Operations Data Analyst (Full-Time)

- Analyzed 20+ derivatives trading transaction data daily. Created **Tableau** dashboards and presented to clients and directors.
- Modeled derivatives product pricing and transactions with the application of Black-Scholes Pricing Models.
- Queried from internal database using **MySQL**, quantified risk metrics and developed KRI reports using v-lookup, pivot table.
- Streamlined internal risk management processes and improved drafting time by 20%.
- Led weekly conference calls with Asian, European teams and collaborated with other business divisions.

PROJECTS (More on Website)

Music Distributed Microservice Application Project   Mar 2022

- Built a music containerized application, including 3 microservices, running within **Kubernetes** cluster on **AWS EKS**.
- Implemented **REST** requests between local machine and **DynamoDB** with functions of reading, creating, editing, deleting.
- Scaled and tested the application by load testing using **Gatling** simulation to observe accuracy, latency, system utilization, and health of system mesh of the microservices under small, medium, and large request load scenarios.
- Visualized and quantified the performance metrics using **Kiali**, **Grafana** dashboard, and **Prometheus**.
- Managed the project using **Scrum** and Kanban in a team with 3 teammates and applied version control using **GitHub**.
- Explored failure scenarios and tested with remediation to inspect multiple system evolving.

YouTube Trending Video Prediction Project   April 2022

- Built a **data pipeline** to retrieve updated trending video data in Python on demand.
- Performed EDA to data and presented interactive analysis results in **Tableau** dashboards.
- Implemented K-Means and hierarchical **clustering models** to analyze keywords.
- Built a data product (a website) to help YouTubers produce more popular videos.