Qianyi Sha in ⌂
Software Engineer
shhh9712@gmail.com (949)-383-0343 Seattle, WA

Profile Summary

- Specializes in data engineering and full-stack development, leverages advanced cloud based platforms like **Snowflake**, **Databricks**, and **BigQuery** to build and optimize scalable data pipelines. Demonstrates strong proficiency in Python, SQL, **PySpark** focusing on custom data software solutions to enhance operational efficiency. Has Great knowledge on data modeling, **ETL/ELT** practices, data pipelining and automation.
- Full-stack developer with a strong command of programming languages including **Python**, **SQL**, **C/C++**, **Java**, and **JavaScript**. Proficient in designing UIs with **React** and **Next.js**, writing server-side code and building **APIs**, and manipulating SQL & NoSQL **databases**. Solid understanding of authentication, security, and version control using **Git** and **CI/CD processes**.
- Proven leadership in project management and technical tool development, effectively leading cross-functional teams to improve software quality and data processing efficiency.

Education

University of Washington, Seattle ⌂
M. Sc. in Electrical & Computer Engineering
Seattle, WA   **GPA: 3.87/4.0**

University of California, Irvine ⌂
B. Sc. in Computer Science & Engineering
Irvine, CA
3 appearances on Dean's honor list

Languages

English: Fluent   **Chinese**: Native level

Technical skills

Languages: Python, SQL, C/C++, Java, JavaScript, HTML, CSS
Data platform/Tools: Snowflake, Databricks, Bigquery, DBT, Azure, AWS, GCP, Postgres, PySpark, Pandas, Redshift, Hadoop, Iceberg, MongoDB, Hive, R, Scala, Go, Airflow, Fivetran, Informatica, Docker, Linux.
Frameworks & Libraries: React.js, Django, Spring, Scikit-Learn, PyTorch, Git, Kubernetes, Tableau, Power BI, Looker.

Professional Experience

Ascend.io ⌂   Field Data Engineer / Solution Architect   **Menlo Park, CA** Oct. 2022 - Feb. 2024

- **Advanced Data Pipeline Architecture**: Led the creation and optimization of scalable data pipelines processing 100M+ rows of record per day across Snowflake, Databricks, BigQuery and connecting with systems like Salesforce, PagerDuty and ServiceNow using SQL, PySpark, and Python. Customized internal tool that dynamically controls the structure and deployment of data pipelines. Eliminate the needs of manual pipeline creation for 50% of the data sources, enhanced data processing efficiency by 40% and ensuring scalability.
- **Tech Lead for Platform Tool Development:** Spearheaded the creation of a platform migration CLI tool in Python, integrating Typer, REST API, and Google Protobuf for efficient data pipeline transfers. This initiative successfully transitioned over 1000 pipelines for five clients to a new platform, also incorporating data quality checks. My role also extended to managing project logistics within an Agile framework, ensuring smooth customer transitions.
- **Platform Development and Maintenance:** Contributing to various data connector frameworks. Ensure platform stability by troubleshooting production issues on Kubernetes clusters using monitoring tools such as OpenTelemetry and Grafana. Identify and resolve problems within our tech stack that includes React, gRPC, Go, among others.
- **CI/CD Process Efficiency**: Implemented efficient CI/CD workflows for data pipeline deployment using CircleCI, Github Actions, Jenkins and Python for different customers, focusing on deployment and testing automation.
- **Solution Architecture and Customer Success**: Directed solution architecture, aligning with client needs and platform strengths, and developed comprehensive customer success strategies to ensure smooth onboarding, product adoption, and cost optimization.

Youyuan ⌂ Software and Data Engineer Intern   **Beijing, China** Jan. 2021 - June. 2021

- **Data Pipeline Optimization**: Enhanced main platform's data processing speed of 100k+ daily records and increased subscription rates by 10%+ using ML models (Decision Trees, Logistic Regression) for user interest analysis.
- **Cross-Team Collaboration Leadership**: Led stakeholder engagement and cross-functional team collaborations, streamlining project execution and removing roadblocks for improved efficiency.
- **Technical Proficiency in Data Solutions**: Utilized Python, SQL, Jupyter, and Scikit-learn for solution development, leveraging senior engineer mentorship for best practices adoption in software and data engineering.

Project Experience

Development of AI Companion Chat App project_link ⌂   Feb. 2024 - onGoing

- Developed an AI chat app that allows users to create and converse with their own AI persona.
- Utilized the **Next.js React** framework along with **Typescript** and **Tailwind** CSS, employed **Postgres** and **Prisma** for database management, and integrated **Stripe** for payment and subscription handling.
- Built AI chat functionality using **Langchain** and **OpenAI**, with **Pinecone** serving as the vector database.