Sriranjan Sribhashyam

ranjan31051997@gmail.com | linkedin/ranjansrinivas/ | github.com/rkaahean | ranjan.dev

Summary

I am a Data Engineer with 4 years of experience in building data-pipelines and analytics engineering using Python along with dbt and Snowflake. I am eligible for a **2 year visa**.

Technical Skills

Languages & Frameworks: SQL, Python, Typescript, , React, Next.js
DevOps: IaC (Terraform), CI/CD, Docker, Airflow, Kafka, AWS SAA Certified (SAAC02)
Data Management: ETL pipelines, dbt, Snowflake, FiveTran, PostgreSQL
Data Science: Tableau, pandas & dask, Py(Spark), OpenAI APIs

Experience

Software Engineer Feb 2021 – June 2024
ChipperCash *Los Angeles, USA*

- Implemented an ELT batch pipeline for computing user risk rating using Python, pandas (dask) & dbt models processing 10M+ rows in ∼2 hours and providing analysts with fresh data in Snowflake.
- Deployed ML models for inference as a microservice using Flask REST APIs, served over an event-driven architecture using RabbitMQ that handled >1000 requests per minute.
- Optimized data loads from Snowflake using Apache Arrow, which sped up pipeline speeds by up-to 10x and decreased credit usage up-to 16x.
- Set up dbt for the analytics function which facilitated higher quality and faster data, for usage in upstream reporting tools like Looker & Tableau.
- Devised an automated Python data pipeline that declined up to 100 chargebacks per minute, thereby saving 200 hours of manual labor and preventing potential losses of ∼$4M.
- Collaborated with PMs and Marketing to build a full-stack in-app messaging system, spanning React Native screens in the Mobile App, to backend REST APIs with TypeScript - thereby saving ∼$80k/year by phasing out vendors.

Data Engineer May 2020 – Dec 2020
Data Sleek *Los Angeles, USA*

- Collaborated with 5+ clients to design, translate and implement business requirements into a technical spec.
- Delivered SQL stored procedures that perform analytical queries to detect and alert customers for changes in desired metrics, saving up to 20% of the base cost per item.
- Performed benchmark tests for several SingleStore configurations to determine optimal configuration for Time Series data, saving $475/year while improving performance by 400%.

Software Engineer Intern May 2018 – July 2018
Samsung Research *Bengaluru, India*

- Set up Open Network Automation Platform on a local server on top of Openstack, using docker containers, laying the foundation for dynamically scalable ML applications based on resource requirements.
- Upgraded the ETL pipeline by deploying to Apache Kafka for streaming data to the microservice.

Education

University of California, Los Angeles Sep 2019 – Dec 2020
Masters of Science in Business Analytics *Los Angeles, USA*
College of Engineering Guindy Jul 2015 – May 2019
B.S. Computer Science & Engineering *Chennai, India*

Projects

Youtube Video Search | *Python, OpenAI, Next.js, PostgreSQL, Typescript* github.com/youtube-search

- Created a website that video-searched podcasts of a popular YouTube channel. Converted raw data into audio and transcribed it using whisper.cpp.
- Used pgvector to store and query embeddings, and generated context summary with gpt-3.5-turbo.

MyAnimeList CLI | *Rust* github.com/mal

- Created an interactive CLI client in Rust to list anime based on a user's profile.
- Display detailed information regarding a show by parsing API.