

Jay Kodeswaran

SENIOR YEAR COMPUTER SCIENCE MAJOR

☎ +1 (206) 307-2853 | ✉ kodesjay@gmail.com | 🏠 jaykodes.me | 🌐 jaykodes | in jaykodes

Skills

Languages Python, SQL, C++, C, Java, Scala, JavaScript, PHP, HTML/CSS

Systems / Libraries BigQuery, Databricks, Snowflake, MySQL, PostgreSQL, MongoDB, Redis, TensorFlow, Pandas, NumPy

Frameworks / Tools Spark, Hadoop, Airflow, Dagster, Kafka, Presto, Hive, Flink, Redshift, S3, Docker, Alembic, DBT, Git

Work Experience

Statsig Inc - Data Engineer

Python, Spark, Databricks, BigQuery, Dagster

Bellevue, WA

May 2023 - Present

- Will create and improve data ingestion flows and data tooling to optimize the internal stats engine

GTS - Software Developer

Python, C++, Airflow, MySQL, Docker

New York, NY

Sept 2022 - Dec 2022

- Scaled the risk system to manage **100+** portfolios by achieving **concurrency** and reducing the runtime of each portfolio by **38%**
- Built reporting infrastructure that notifies profit discrepancies, increasing discoverability by **10%** and saving **\$2000+** per month
- Redesigned the database schemas to follow **3NF**, alleviating **14%** of space usage and optimizing data pipeline runtimes by **6%**
- Enhanced the asset classification system by accommodating for **70+** OTC securities improving the value of risk accuracy by **2%**

Wish - Data Engineer

Python, Airflow, Spark, Hive, Presto

San Francisco, CA

Jan 2022 - Apr 2022

- Engaged with data scientists to develop data models and pipelines that discovers up to **\$250,000** of chargeback fraud per week
- Developed a ML pipeline that trains the model on **5+ GB**, reducing training time by **15%+** and increasing model accuracy by **4%**
- Created and maintained **20+** data pipelines, migrating **30+ GB** daily across different infrastructures (**MongoDB, Redshift, S3**, etc)
- Parallelized data flow, created partition schemes, and utilized optimal data engines to decrease the runtime of ETL pipelines by **23%**

Equitable Life of Canada - Software Developer

Python, PostgreSQL, Flask, Pandas, NumPy

Waterloo, ON

May 2021 - Aug 2021

- Spearheaded the construction of a scalable batch processing system, increasing data-collection efficiencies **5x** and saving **\$24,000**
- Automated ETL processes across **millions** of rows of data, reducing manual workload by **18%** and processing times by **22%**
- Produced a web application that queries **200+** stocks and displays trade volatility, improving trading detection algorithms by **3%**

Genellipse Inc - Data Analyst

Python, MySQL, Pandas, NumPy, TensorFlow

Mississauga, ON

Sept 2020 - Apr 2021

- Increased recruitment effectiveness by **15%** by developing a **deep neural network** that classifies members with **91%** accuracy
- Pioneered a reporting pipeline that transforms raw payroll data while maintaining **98%** uptime, saving **4+** hours a week
- Performed cluster analysis on member spending to create personalized marketing campaigns, increasing weekly revenue by **\$150+**

Programming and Design Projects

WatDFS - Systems Design Project

C++

github.com/jaykodes/WatDFS

- Built a **distributed file system** that utilizes an **upload/download model**, updates files using **freshness checks**, and uses **RPCs**

Virtual Voice Assistant - Full Stack Project

Python, Javascript, HTML/CSS, Django, scikit-learn

github.com/jaykodes/Alfred

- Developed a desktop assistant that uses **natural language processing** and **logistic regression** to identify and fulfill user needs

Education

University of Waterloo

Candidate for Bachelor of Computer Science (3.7 / 4.0)

Waterloo, ON

Sept 2019 - Apr 2024

- **Relevant Courses:** Machine Learning, Distributed Systems, Database Systems, Algorithms, Data Structures, Operating Systems