Prithaj Nath

⚲ Burlington, VT   ✉ prithajnath@gmail.com   📞 (518) 645-1487   ⧉ www.prithaj.dev

▦ iamprithaj   ⚙ prithajnath   ⸙ prithajnath

Summary

Software Engineer with 7+ years of experience building scalable ETL pipelines, deploying machine learning models, and optimizing data infrastructure in cloud environments. Skilled in Python, SQL, Docker, and AWS with a strong academic focus on complex systems and applied machine learning.

Technologies

Programming & Scripting: Python, SQL, Bash, Ruby, JavaScript

ML/AI & Data Science: Pandas, NumPy, scikit-learn, matplotlib, seaborn

Cloud & DevOps: Docker, Terraform, Kubernetes, AWS (EC2, Lambda, ELB, EKS, SNS, SQS)

Data Infrastructure: PostgreSQL, BigQuery, Snowflake, Redshift

Education

University Of Vermont   *Aug 2024 – May 2026*
M.S. in Computer Science with a specialization in complex systems

SUNY Plattsburgh   *Aug 2013 – May 2017*
B.S. in Computer Science

Projects

Analysis Of Superuser Churn On StackOverflow (https://bit.ly/41uz0yI)   *Aug 2024 – Dec 2024*
- Set up ETL pipelines using Python, SQL, and Bash to transfer 90+ GB of Stack Overflow user data in XML to BigQuery
- Used BigQuery SQL to identify a superuser cohort (~22k users) and generated time-series data by calculating their rolling six-month averages of posts
- Detected churn patterns by fitting exponential decay models on user activity time-series

Experience

Software Engineer   *New York, New York*
Singuli, Inc   *May 2022 – May 2024*
- Reduced median client onboarding time by 80% by implementing ETL pipelines to daily ingest sales, inventory, and revenue data using Apache Airflow, Python, and SQL
- Improved forecast quality by 20% by adding data sanity checks for machine learning models (k-means clustering, time-series regression)
- Increased number of prospective clients by 30% by implementing a competitor dashboard analyzing inventory, revenue, and sales data from client competitors using Flask, SQL, PostgreSQL, and Redshift
- Improved API response time by 50% (10s to 4.8s) through PostgreSQL query optimization enabling real-time reports on historical sales and inventory data

Integration Engineer   *Burlington, VT*
Faraday, Inc   *June 2019 – May 2022*
- Implemented ETL pipelines on Kubernetes (EKS) to ingest customer and geospatial data from 50+ sources using Python, Ruby, SQL, and BigQuery
- Built micro-services in Python and Ruby to seed Facebook lookalike audiences with AI-powered cohorts, increasing client ROAS by 50%

- Implemented a micro-service in Python to deploy AI-powered bid modifiers to Google Ads at the zip code level, increasing ad campaign conversion rates by 38%

Jr Data Scientist *Burlington, VT*

IXIS Digital *Aug 2017 – Sept 2019*

- Analyzed e-commerce data from Google Analytics and Adobe Analytics, generating ad-hoc reports using R, SQL, and Python
- Implemented CI/CD pipelines for bare metal servers using Docker and GitLab CI
- Built a Python tool to concurrently scrape 100+ sites and track changes in critical DOM elements