Dante Mazza

dantemazza.netlify.app
github.com/dantemazza
+1 425-542-4364
dantemazza1@gmail.com

Skills

Languages: **Python, Java, Scala, Kotlin, C, C++, SQL**

Libraries/Frameworks: **Flask, Spring Boot, Spark, Hadoop, TensorFlow, Keras, Delta Lake**

Tools/Infrastructure: **AWS, Docker, Kubernetes, Prometheus, Kafka, Linux, Terraform, Redis, Nginx**

Work Experience

Splunk   *Toronto, Canada*

Software Engineer II   *July 2023 - Present*

- Developing scalable and performant (50 TB/day) AWS-hosted Kubernetes+Java/Spring microservices with PostgreSQL/RDS, SQS for messaging, Prometheus for monitoring, Splunk, GitLab CI/CD, Terraform
- Built and maintained Spark/Scala batch pipelines using Delta Lake on S3, AWS EMR, and Step Functions, to process 1PB+ of logs (VPN, Windows, cloud storage), detecting anomalies across 30M+ users and devices daily
- Designed and implemented an E2E testing framework for ML models with Python, AWS, and GitLab CI/CD. Successfully advocated for its adoption, reducing testing and deployment time by 98% - from weeks to hours
- Achieved a 30% reduction in team AWS costs, saving 150k per year, by optimizing Spark/Scala ETL pipeline (AWS Glue, Kinesis, Delta Lake) and downstream Spark ML jobs on EMR with performance tuning and code optimization
- Reduced mean time to resolution by owning production debugging across distributed services; consistently identified root causes, improved alert quality, and produced more actionable monitoring
- Led implementation of control plane for security detections that increased customer content offering by 2x

Amazon Keyspaces (AWS)   *Bellevue, WA*

Software Development Engineer Intern   *Aug 2022 - Dec 2022*

- Authored a transactional query service using Kotlin and internal AWS services for a managed Cassandra product
- Researched open-source Cassandra implementation and functionality for batch queries and designed for parity

Splunk   *Toronto, Canada*

Software Engineer Intern   *Jan 2022 - Apr 2022*

- Prototyped multi-tenant Spark job execution with AWS (EMR, IAM, STS, Step Functions, Lambda), Terraform, Python, Docker, Java, Spring
- Automated tenant provisioning with shell+GitLab/internal APIs, saving 6 hours of onboarding per customer

KCM Solutions   *Toronto, Canada*

Software Engineer Intern   *Jan 2021 - Apr 2021*

- Built Node.js and Python serverless backend services for a chatbot with various APIs (Twilio, MS Graph, Watson)
- Created containerized CentOS enterprise software environments with Docker/Compose for centralized use; established network communication, LDAP authentication, Postgres connections

Wisedocs   *Toronto, Canada*

Machine Learning Engineer   *May 2020 - Sept 2021*

- Implemented and optimized machine learning models, including CNNs (TensorFlow, Keras), NLP models (Hugging Face Transformers), and k-means clustering (scikit-learn), achieving a 2x improvement in prediction accuracy
- Developed document clustering algorithms in Python, reducing prediction time by 85%
- Determined pipeline evaluation metrics with SQL queries and Pandas, providing ML performance benchmarks
- Developed a training method for pre-serialized datasets via Keras' batch interface; reduced training time by 90%

Education

University of Waterloo | BASc, Computer Engineering   *Sept 2018 - May 2023*

- Deans Honour's List (first year)
- Relevant coursework: Distributed Systems, Compilers, Operating Systems, Networks (C, C++, Java, Rust)