

Alexandre CRUEL

Senior ML Engineer, M.Sc.

170 rue Rioux, H3C 0X3, Montreal
+1 (514) 600-6858 / +33 667928245
alexandre.cruel@orange.fr



After 3+ years as a Data Engineer in an Artificial Intelligence team, I wish to continue my AI and data journey to develop, productionize and maintain data pipelines and ML solutions on the cloud.

Language: French: mother tong, English: bilingual, Spanish: basic

EXPERIENCES

June 2020 – now, Senior ML Engineer in the AI Platform - National Bank of Canada

- Responsible for a team of data engineers and interns
- Work closely with partners from various business units and take care of the development and deployment of Artificial Intelligence components in production.
- Develop, manage, and maintain high-performance data ingestion mechanisms (ETL & ELT using Apache Spark and Apache Beam) to feed various machine learning applications integrated into live business processes (daily/monthly processes, near real-time use cases hosted on GCP and AWS)
- Creation of a data quality tool to monitor ingested data and display data quality score results in a UI based on the 6 data quality pillars (TypeScript, Angular)
- Fully participate in the analysis, design, and implementation of business performance approaches, developing tailored solutions and creating positive impact and lasting results.
- Collaborate with the Software Engineering Lead, data source management teams and DevOps evangelism team in coaching best practices in artificial intelligence development.
- Participate in studies on the state of the art of data ingestion software design for artificial intelligence applications in python deployed on GCP or AWS.
- Supervise development practices upstream from the prototyping phases and take charge of the design of the product to be deployed downstream. Making sure the proposed architecture is created to last and versatile (working on Data Mesh architecture).
- Support the delivery of data management functionality for AI solutions.
- Writing of operational models for the management of errors on predictions generated in production.

September 2019 – June 2020, Machine Learning Engineer in AI Factory team - National Bank of Canada

- Team Scrum Master: Responsible for managing scrum daily meetings and ceremonies
Integrate multidisciplinary teams in collaboration with software engineers and machine learning experts. Collaborate in building an integrated process for developing, presenting, prioritizing, and leading the implementation of AI solutions.
- Building Data pipelines on Google Cloud Platform for AI models:
Management of BigQuery Data Warehouse, Cloud Dataflow and Dataproc data pipelines.
Working daily with Google Cloud Storage, Google Kubernetes Engine, Pub/Sub.
- Productionizing AI models by developing reusable libraries for AI purposes
- Managing data and dev environments for AI scientists on GCP:
GCE (Data Tokenization & Encryption, Security measures implementation, Identity and Access Management...)

April 2019 – September 2019, Artificial Intelligence Intern - National Bank of Canada

Operationalization of artificial intelligence models for a banking platform

- Development of a reusable Flask API for Machine Learning models which exposes the different endpoints of the model (preprocessing, train/test, predict_batch)

- Implementing a Data Ingestion process to fetch data from on premises data sources to the Google Cloud Platform ecosystem using Cloud Storage, BigQuery and Cloud Dataproc with the capacity to treat various types of Input/Output (CSV, JSON, Parquet, JDBC, Avro). The solution was also designed to perform aggregations and tokenization on the extracted data.
- Creation of a training platform for AI scientists to promote data and software engineering best practices.
- Working in an Agile environment

September 2018 – April 2019, Data Scientist - Swapcard

1st place at ESME symposium, 2019 (Information Systems category)

Recommendation system for a trade exhibitions phone app (Master's end of studies project)

- Python and SQL based solution
- Use of dbScanClustering, Word2Vec (Google), nltk, fastText, Docker technologies
- State of the Art research to handle cold start problem (using stereotype method)
- Content based and Collaborative filtering

2017 – 2019, Mathematics teacher - Acadomia

- Tutoring for high school students with difficulties

SKILLS

Software Development / Data Engineering:

- Cloud:
 - GCP (BigQuery, Dataflow, PubSub, GKE, GCS, AI Platform and more)
 - AWS (S3, ECS, Step Function, SQS/SNS, Glue, DynamoDB, Lambda, Redshift)
- Python, C++, JAVA, JavaScript (NodeJS, Angular), R, Web
- Apache Spark, Apache Beam
- SQL: PostgreSQL, MySQL
- Microservices: Docker, Kubernetes
- CI/CD: Jenkins, Terraform, circle.ci
- DevOps: Datadog, Splunk, SonarQube, NexusIQ, Veracode
- Architectural diagrams on Draw.io

Professional Certifications:

- Coursera: [Specialization Professional GCP Data Engineer with google cloud](#)
- Udemy: [AWS essentials](#)
- Udemy: [Python 2020](#)
- MILA: [6th MILA/IVADO Deep Learning School](#)

EDUCATION

2014-2019: Master's degree in Big Data and Information System | ESME Sudria Engineering school, Paris

2017: International Semester | California State University of Monterey Bay, CA

2014: Scientific high school diploma with honors | Lycée Berthollet

INTEREST

Saxophone Alto, Video editing, Model construction, Mountain sports, Running, Basic plane pilot license