

*Double Masters of computer engineering, dedicated **Data Analyst & Data Engineer** with 6 years' experience in **web-based front end development**, **backend pipeline design** and **Datawarehouse implementation on premise / clouds***

## Qualifications

- Extensive knowledge in **Data Modeling**, **Data pipeline** and **Data Visualization** design
- Rich experience with **AWS** and **Azure** cloud implementation of BI deliveries
- Hands on **Machine Learning** Assignments using **neural network**, liner regression, classification, k-means
- **Web-based** application front-end design, and back-end development
- **3-year** experience in Data **ETL/ELT** using SQL, Stored Procedure, Python, SparkSQL, PySpark, Linux shell script
- **2-year Dashboard design** using Power BI and Tableau
- **2-year** experience in **DEVOPS CICD**, including Process Design, Development of release pipeline, yaml and template on cloud platform
- Great understanding of project management in **Agile** and waterfall methodology, familiar with Jira, Confluence
- Strong problem solving skills and creative trouble-shooter, willing to help others, a true team player

## Education

Master of Engineering, Electrical & <b>Computer</b> Engineering   University of Ottawa	09/2020 - 06/2022
Master of Engineering, <b>Software</b> Engineering   Western University, Toronto	09/2019 - 10/2020
Bachelor, <b>Computer</b> Engineering   Shandong University of Finance & Economics, China	09/2015 - 06/2019

## Experience

**Data Analyst | RBC Royal Bank of Canada, Toronto** 05/2021 - 12/2021

- Developed Global EDI Data Exchange Framework between SAP SAF and Vendor-Endpoints of RBC through REST API, SFTP, including format conversion from iDoc to CSV, Naming Convention and Orchestration for 30 workflows, which covers procurement, employee, invoicing etc
- Re-designed the Web UI, Data Recording and BI ETL process of RBC Cost Redistribution Platform to better allocate RBC property/transit costs in terms of space management, financial calculation and payment allocation
- Containerized in-house web applications with CI/CD development and process using Jenkins and Introduced automated way to deploy applications to Enterprise OpenShift Cloud Platform across environments
- Created Workforce Management Portal for managers maintaining employee vaccination status in FutureWork database which provides information to VPs for company's work mode decision making
- Developed & integrated an Enterprise Application Portfolios Management Dashboard in RBC IIPM system to provide an overall picture of IT Management and Governance for higher management's decision making
- Corporate Real Estate Information Database investigation and housekeeping base on source identification, naming convention, last update time and usage analyses

*Tools: OpenShift Cloud Platform, Azure SQL DB, Procedure, Python, Plotly Dash Framework, Jenkins, Docker, yaml, Nexus, Jfrog*

**Digital Market Analyst | Trip.com Group, China** 06/2020 – 09/2020

- Developed an online analytical plugin on monitoring travel products using Python, increased certain product sales by 15% through KPI analysis, such as product clicks, customer viewing time
- Designed A/B testing to monitor and record the KPIs, such as time-on-page, click-through rate and conversion rate
- Consolidated website performance tagging structure, created analytical requirement documents for travel products
- Conducted conversion latency analysis and created latency projection model which provided key insights to the customer value calculation and attribution.
- Created sustainable automated solution for report scheduling and delivery to ensure cross-department reference.
- Automated project procedure in Python and VBA and save 90% time consumed.

*Tools: Python, HTML5, CSS3, Angular.js, React.js, NodeJS, JavaScript, Access, Excel*

## Cloud Projects

### Project 1 – Azure Databricks for bank data ingestion & Transformation 04/2020 - 05/2020

- Coded ingestion pipelines that load csv, json files into Databricks, using sparkSQL, pyspark and python
- Created DB, Tables, Views and transform the data in Delta lake
- Implemented full refresh and incremental load patterns using partitions
- Created Azure data factory pipelines to execute Databricks notebooks and ADF triggers for orchestration

*Tools: Azure Databricks, Azure Data Factory, Azure SQL DB, sparkSQL, pySpark, Python, ADLS Gen2, Storage Account, snowflake*

### Project 2 – AWS End to End ETL Pipeline Design in python 03/2020 - 04/2020

- Developed AWS S3 data pipeline for security trading summary report using python framework
- Designed ETL in A Functional and object oriented programming way to extract, transform and load from AWS S3 source and target with best practice
- Used a meta file for job control and identified deployment process across environments for production ready

*Tools: AWS S3 buckets, Python 3.9, Jupyter Notebook, Github, Visual Studio Code, Docker Hub and the boto3, pyyaml, awscli, jupyter, pylint, moto and the memory-profiler*

## Academic Project Experience

### Project 1 – AWS Image Resizing App 04/2020 - 05/2020

- Deployed an image processing app using Elastic Beanstalk service & EC2 instance in AWS Cloud
- Deployed an image resizing app with a lambda function that gets triggered whenever an image is uploaded to an S3 bucket storage, then it creates a smaller version of that image into a destination bucket

### Project 2 – BI Solution End to End Implementation 03/2020 - 04/2020

- Performed reseller sales data profiling, gap analysis in Customer Sales Database with complex Ad-hoc SQL
- Designed data mart star schema & multidimensional model and create DB, Schemas and tables in Database
- Developed TSQL Stored procedures, SSIS service implementing Data ETL
- Created SSAS analysis to predict sales amount and reseller location using regression and classification algorithms

### Project 3 – Facial Key-Point Recognition & Emotion Detection 11/2019 - 12/2019

- Developed a facial key-point recognition & emotion detection project using Neural Networks, and optimized tuning process, to detect eyes, nose and mouth positions. Reached prediction accuracy of 81%
- Implemented emotion detection based on facial key points, and designed further applications like locate human face, generate average face, and face swap

### Project 4 – Music Review Website 10/2019 - 11/2019

- Created a music review website similar to Spotify in Full Stack, using HTML5, CSS3, Angular.js, React.js in Front-End, and NodeJS, JavaScript in Back-End
- Implemented functionalities such as user register/login, third-party login, search/add/review songs, manage playlist, site user interaction, site manager

## Skills

Program Language	Python, Java, SQL, Linux Shell Script
Database	Azure SQL DB, Snowflake, SQL Server
ETL	SSIS, Procedure, Python Framework, Pyspark, SQLSpark
Dashboard	Tableau, Power BI, Google Analytics
Web Development	HTML5, JavaScript, React.js, Angular.js
Cloud	Docker, RedHat OpenShift, Kubernetes, AWS, Azure Datalake, ADF, Databricks
DevOps	Jenkins, Docker, yaml, Nexus, Jfrog, yaml, GITHUB
Project Management	Jira, Confluence, draw.io, ServiceNow