

# Marc Cheng



438 886 8425 | [marccheng@outlook.com](mailto:marccheng@outlook.com)  
Mathematics Specialist | Data Scientist

## SKILLS

### ■ Machine Learning

pytorch, tensorflow,  
keras, scikit-learn

### ■ Data Science

pandas, numpy, scipy,  
matplotlib, seaborn,  
SAS, Tableau, Power BI

### ■ SaaS

Amazon AWS, MS Azure,  
Google Colab, BigQuery,  
SAP Hybris, IBM DB2,  
Informatica P360

### ■ Programming

Python, Jupyter Notebook,  
R, SQL, HTML, CSS,  
JavaScript, Node.js

### ■ Shell Scripting

Bash, CMD, PowerShell

### ■ Languages

English, French, Mandarin

### ■ Interest & Hobby

Skateboarding, Rock  
Climbing, Piano, Guitar,  
Solving 3x3 Rubik's cube  
under one minute

## AWARDS

### ■ University of Toronto

2023, 2022, 2021 | Dean's List

### ■ L'Université Laval

Concours Opti-math  
2015 | 5<sup>th</sup> place in Province  
2014 | 7<sup>th</sup> place in Province

### ■ University of Waterloo

Gauss Mathematics Contests  
2013 | Distinction

## ACADEMIC

### Multivariate Regression Analysis on Mortality Rate University of Toronto

Independently designed  
univariate and multivariate  
models to validate assumptions,  
using the Annual Health Survey  
(AHS) data from 2012 to 2013.

## PROFESSIONAL EXPERIENCE

### Rexel Canada | June 2023 – Present

#### Product Data Specialist | Mississauga, ON

- Successfully completed a year-long implementation project of the new PIM system, developed key functionalities in Informatica Product 360.
- Carried out the entire end-to-end (ERP - PIM - SAP) testing process in the UAT environment, in collaboration with the TIBCO BW5 team.
- Handled daily product quality maintenance of 200,000 unique SKUs, resolved JIRA tickets using Python scripts and custom SQL queries on IBM AS400.
- Enhanced product data quality across divisions by collaborating with industry-standard data sources and designing automated daily data quality rules.
- Developed dataflow control processes using AWS CLI and PowerShell scripts.
- Participated in RexelGPT, a now live LLM project during its beta testing phase.

**Key Skills:** PIM / MDM Software, ERP Software, SAP Hybris, AWS, Python, SQL

### Deloitte Canada | July 2021 – April 2022

#### Data Analyst (Co-op) | Toronto, ON

- Validated anti-money laundering models written in R and Python for Scotiabank.
- Completed the validation of 8 AML models using both supervised and unsupervised clustering techniques, including KNN and K-Means.
- Extracted 1,000,000 to 20,000,000 entries of transactional data from an Oracle database using SQL and verified data integrity with Jupyter Notebook.
- Validated redemption models for LoyaltyOne using SQL Server Management Studio in conjunction with SAS, providing insights into customer conversion rates for Air Miles and American Express.
- Collaborated on the valuation of 20 derivative contracts, including interest rate swaps and commodity swaps (e.g., natural gas and crude oil), using Bloomberg Terminal.

**Key Skills:** R, Python, SQL, SAS, Jupyter Notebook, Bloomberg Terminal

## PROJECTS

### Implementation of GPT-2 Large Language Model

- Designed the model with 768-dimensional embeddings and 12-head attention.
- Trained with 162M parameters, applied weight masking to prevent overfitting.

**Key Skills:** Architecture Design, Natural Language Processing

### Deep Residual Networks for CIFAR-10 and CIFAR-100 Dataset

- Implemented 34-layer and 101-layer residual networks for CIFAR-10 and CIFAR-100.
- Accelerated the training process by enabling CUDA on an RTX 3060 (12GB).

**Key Skills:** Deep Learning, Computer Vision

### Quantitative Analysis in Cryptocurrency Trading

- Applied unsupervised clustering technique (DBSCAN) to identify outliers.
- Used RNN and LSTM in conjunction with fitting techniques to predict future prices.

**Key Skills:** Clustering, Time Series Analysis, Regression

## EDUCATION

### University of Toronto | September 2019 – April 2023

#### H.B.Sc. Honours Bachelor of Science | Toronto, ON

- Mathematics Specialist (Focus in Probability/Statistics) | GPA 3.6/4.0