

Zhongmang (Marc) Cheng

Data Engineer · ML Systems | Right to Work: TTPS Visa

Canada: +1 438-886-8425 | Mainland China: +86 156-5240-1103 | marccheng@outlook.com
chengmarc.com | linkedin.com/in/chengmarc | github.com/chengmarc

EDUCATION

University of Toronto

Toronto, ON

Honours Bachelor of Science — Mathematics & Its Applications

Sep. 2019 – Jun. 2023

- Specialist in Probability and Statistics; Major in Mathematics
- GPA: 3.6/4.0, High Distinction | Dean's List: 2021, 2022, 2023
- Relevant coursework: Probability Theory, Statistics, Real Analysis, Methods of Data Analysis, Time Series Analysis, Multivariate Data Analysis, Nonlinear Optimization, Intro to Computer Science

PROFESSIONAL EXPERIENCE

Product Data Specialist

Jun. 2023 – May 2025

Rexel Canada

Mississauga, ON (On-Site)

- Led end-to-end implementation of a new Product Information Management (PIM) system, defining data models and core workflows in Informatica Product 360 to support enterprise-wide product operations.
- Acted as the primary data interface between SAP Hybris, ERP (IBM AS400), and external industry data sources, ensuring consistency across 200,000+ SKUs.
- Designed and executed UAT frameworks and validation rules, coordinating cross-functional testing with offshore engineering teams (India, France) to ensure production readiness.
- Designed and implemented automated ETL data quality control pipelines, translating business requirements into daily validation rules.
- Developed dataflow orchestration and CI/CD control scripts using AWS CLI and Bash, enabling automated batch pipeline deployment and monitoring.
- Contributed to enterprise LLM initiative (RexelGPT), focusing on system performance optimization and real-world deployment constraints during beta testing.

Data Analyst (Co-op)

Jul. 2021 – Apr. 2022

Deloitte

Toronto, ON (Hybrid)

- Performed validation of enterprise AML models for a Tier-1 Canadian bank, assessing model assumptions, feature selection, and statistical robustness.
- Evaluated supervised and unsupervised risk models (KNN, K-Means, XGBoost) using large-scale transactional datasets (1M–20M records), ensuring regulatory compliance with model risk standards (SR 11-7 framework).
- Conducted model logic and data pipeline verification using SQL, Python, and R to identify inconsistencies between theoretical design and production behavior.
- Supported valuation and validation of derivative pricing models, translating financial contract structures into quantitative assumptions using Bloomberg Terminal.

FREELANCE EXPERIENCE

Front-End Developer

Jun. 2025 – Present

DoubleTrends

Toronto, ON (Remote)

- Designed and developed doubletrends.com from scratch using Tailwind CSS and Node.js, delivering a fully responsive interface for desktop and mobile.
- Managed deployment on Cloudflare Pages; integrated Stripe payment processing to support gated features and user transactions.

ML Infrastructure Contributor

Feb. 2025 – Mar. 2025

Stoic AI

New York, NY (Remote)

- Implemented a MAMBA state-space model (selective SSM with dual-branch projection, depthwise convolution, SiLU gating) for sequence-to-sequence forecasting on multi-feature time series. A de-sensitized version is available on [GitHub](https://github.com).
- Designed a detrending pipeline: logarithmic trend extraction via nonlinear least squares, residual forecasting, and inverse reconstruction.
- Benchmarked against five baselines: SegRNN, LSTM, Seq2Seq LSTM, Attention LSTM, and Transformer.
- Set up cloud-based GPU training and experimentation environments on RunPod utilizing H100 instances to accelerate model iteration.

PROJECTS

- Large Language Model (GPT-2 Style)** | *PyTorch, tiktoken* | [GitHub](#) Dec. 2024
- Implemented a GPT-2-style transformer with 768-dim embeddings and 12-head self-attention.
 - Trained 162M parameters on a custom corpus using AdamW optimizer; tokenization via BPE (tiktoken).
- ERC-20 Memecoin** | *Solidity, Ethereum* | [GitHub](#) Dec. 2024
- Implemented and deployed an ERC-20 token contract to Ethereum mainnet; contract verified on-chain.
- Image Classification** | *PyTorch* | [GitHub](#) Nov. 2024
- Implemented ConvNet on MNIST/EMNIST; ResNet-34 and ResNet-101 on CIFAR-10/100.
 - Compared basic blocks (two 3×3 convolutions) vs. bottleneck blocks ($1 \times 1 \rightarrow 3 \times 3 \rightarrow 1 \times 1$); analyzed interaction between identity shortcut connections and network depth across datasets of varying difficulty.
- Quantitative Analysis Tools** | *Python, Binance REST API* | [GitHub](#) 2023 – 2024
- Built a multi-module crypto quantitative framework: Monte Carlo simulation across 28,000 coins, live trading via Binance REST API, volatility-normalized DCA strategy, BTC halving-cycle price simulation, and on-chain UTXO indicator analysis.
- Market Data Crawlers** | *Python, Selenium, BeautifulSoup4* | [GitHub](#) 2023
- Built three independent data collection pipelines targeting TradingView, CoinGecko, and CoinMarketCap to feed downstream quantitative analysis workflows.

TECHNICAL SKILLS

Languages: Python, R, SAS, Bash, PowerShell

ML & Deep Learning: PyTorch, TensorFlow, Keras, scikit-learn

Data Engineering: pandas, numpy, scipy, Oracle, MySQL, IBM DB2 (AS400), SQL Server, BigQuery

Cloud & Infrastructure: AWS, Azure, RunPod, Google Colab

Front-End: HTML, CSS, JavaScript, Tailwind CSS, Node.js

Visualization & BI: matplotlib, seaborn, Tableau, Power BI

Enterprise & Other: Informatica Product 360, SAP Hybris, Bloomberg Terminal, Git

Spoken Languages: English (bilingual proficiency), French (professional working proficiency), Mandarin (native)

PUBLICATIONS

Zhongmang (Marc) Cheng. “On the Impossibility of Complete Interpretability in Neural Networks: A Measure-Theoretic Analysis.” *Zenodo preprint*, Feb. 2026. DOI: [10.5281/zenodo.18678067](https://doi.org/10.5281/zenodo.18678067)

AWARDS

University of Toronto Dean’s List: 2021, 2022, 2023

Concours Opti-math, L’Université Laval: 5th Place Provincial (2015), 7th Place Provincial (2014)

University of Waterloo Gauss Mathematics Contest: Distinction (2013)