# **\$\square\$ 226 [seven] [nine] [two] 8595 \$\square\$ m375chen@uwaterloo.ca**

# **Mark Chen**

github.com/marko-polo-cheno in linkedin.com/in/mark375chen

#### **WORK EXPERIENCE**

#### Machine Learning Engineer | Kolena Inc

San Fransisco, CA | Jan 2023 - Apr 2023

 TBD - Enhance a rigorous machine learning testing platform to surface hidden behaviors and failure modes with faster automated testing workflows

### Data Scientist | GoldSpot Discoveries Inc

Toronto, ON | Jan 2022 - Apr 2022

- Python, PyTorch, Sklearn, AWS, EC2, S3, Lambda, OpenCV, MLFlow, Streamlit
- Created a multi-class object detector with distributed training on 8 GPUs using the YOLOv5 framework to identify handwritten numerical text on wood/rock textures with 85% precision and 79% recall
- Spearheaded the creation of a spectrometer based on valuable elements by training SimCLR models to cluster images of rock and mineral textures using t-SNE
- Deployed a manager for virtual cloud clusters to sort and predict depths for thousands of miles of rock and drill hole images

#### Data Scientist | eCompliance

Toronto, ON | May 2021 - Aug 2021

- Python, Keras, Spark, Scala, Airflow, Kafka, Looker, SQL
- Created text classification models using LSTM with 97% accuracy on 50 million records to label accident reports
- Applied feature engineering to improve the MSE of a time series forecasting model (FB Prophet) by 15% to predict upcoming trends and types of workplace accidents
- Designed efficient pipelines for incoming user data to be translated if needed, cleaned, and labelled by NLP models

# Software Engineer | eCompliance

Toronto, ON | Jan 2021 - Apr 2021

- Automated entire test suites with Cypress and Flutter for web/mobile platforms to significantly improve DevOps processes
- Identified and resolved multiple bugs using Dart, taking part in a substantial reduction from 100 issues to 20 issues

# QA Analyst and ETL Developer | Teranet

Missisauga, ON | May 2020 - Aug 2020

- Engineered memory efficient ETL packages with advanced SQL validation methods to ensure correctness for half a billion records
- Designed creative solutions in C#, .NET, and T-SQL to standardize complex and ambiguously structured inputs and data

#### **EXTRACURRICULAR ACTIVITIES**

#### **Developer Relations Engineer | Aipaca**

Toronto, ON | Jan 2022 - Aug 2022

- Developer for a toolbox using TensorFlow to predict training times/costs of machine learning models on AWS EC2 instances
- Organized a summit for 150+ participants from multiple clubs and companies to network and deploy ML projects with Aipaca

# Technical Lead + Executive | University of Waterloo Computer Science Club

Waterloo, ON | May 2021 - Sep 2022

Delegated all development for a Discord bot on GCP with a coffee chat, currency, and anti-spam system for 2500+ members

### Open Source Contributor | Microsoft Research Asia

Beijing, China | May 2021 - Mar 2022

• Refined Python code scripts and translated modules involving fundamental AI knowledge from Chinese into English for ai-edu

#### Hackathon Lead Organizer + Mentor | SpartaHacks

Guelph, ON | Jan 2019 - Sep 2019

- Planned event and attained major sponsorships, ensuring event success at no-cost for all 120+ participants
- Ran Blockchain, Python, and Docker workshops

# **EDUCATION**

#### University of Waterloo

Waterloo, ON | Sep 2019 - Apr 2024

- Candidate for Bachelor of Computer Science with Artificial Intelligence Specialization
- Cumulative GPA: 3.7/4.0 or 84%
- Positions: Undergraduate Academic Plans Committee Member Computer Science Club Executive
- Scholarships: President's Scholarship of Distinction Faculty of Mathematics Scholarship

#### **SKILLS**

Languages: Python • SQL • C++ • C • TypeScript • Java • Kotlin • Scheme • Bash • R

Frameworks/Tools: Scikit-learn • TensorFlow • Pandas • Numpy • PyTorch • CUDA • MLFlow • OpenCV • MongoDB • AWS • EC2 •

Databricks • Anaconda • ETL • Node.js • Git • Docker