Mark Chen

📞 226 seven-none-too 8595
✉ m375chen@uwaterloo.ca

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

SKILLS

Languages: Python • SQL • C++ • C • TypeScript • Java • Kotlin • Scheme • Bash • R
Frameworks/Tools: Scikit-learn • TensorFlow • Keras • Pandas • Numpy • PyTorch • CUDA • MLFlow • OpenCV • MongoDB • AWS • EC2 • Databricks • Anaconda • ETL • Node.js • Git • Docker

EDUCATION

University of Waterloo   **September 2019 - April 2024 | Waterloo, ON**
- Candidate for **Bachelor of Computer Science** with Artificial Intelligence Specialization
- Cumulative GPA: 3.7/4.0 or 84%
- Scholarships: President's Scholarship of Distinction • Faculty of Mathematics Scholarship
- Relevant Courses: Artificial Intelligence • Algorithms • Data Structures • Numerical Computation

WORK EXPERIENCE

Data Scientist | GoldSpot Discoveries Inc   **January 2022 - April 2022 | Toronto, ON**
- *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 rock surfaces and wooden textures with 85% precision and 79% recall
- Deployed an automated service that starts and ends a virtual EC2 cloud cluster, for sorting and predicting depths for thousands of kilometers of rock and drill hole images
- 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 after applying dimensionality reduction techniques such as PCA and ICA

Data Scientist | eCompliance   **May 2021 - August 2021 | Toronto, ON**
- *Python, Keras, Spark, Scala, Airflow, Kafka, Looker, SQL*
- Created and trained text classification models using LSTM with 97% accuracy on 50 million records to simplify long workplace accident descriptions into one or two custom labels
- Applied feature engineering to improve the MSE of a time series forecasting model (FB Prophet) by 15% to predict upcoming trends in workplace accidents and potential types of accidents
- Designed efficient pipelines for incoming data to be translated, cleaned, labelled by NLP models, and stored every day

Software Engineer | eCompliance   **January 2021 - April 2021 | Toronto, ON**
- *JavaScript, Dart, Flutter, Cypress, GitHub Actions, YAML*
- Automated entire test suites on web and mobile platforms to significantly improve QA and development processes
- Identified and resolved multiple bugs, taking part in a substantial reduction from 100 issues to 20 issues

QA Analyst and ETL Developer | Teranet   **May 2020 - August 2020 | Missisauga, ON**
- *C#, .NET, SSIS, MS SQL, T-SQL*
- Engineered memory efficient ETL packages with advanced SQL validation to ensure data correctness for half a billion records
- Designed creative solutions in C# to standardize complex and ambiguously structured inputs and data

EXTRACURRICULAR ACTIVITIES

Developer Relations Engineer | Aipaca   **January 2022 - August 2022 | Toronto, ON**
- *Python, TensorFlow, Keras, JavaScript, AWS, SageMaker*
- Contributor to an open-source toolbox for predicting the training time and cost of machine learning models on AWS 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   **May 2021 - September 2022 | Waterloo, ON**
- *Typescript, SQLite, Discord.js, GCP, Docker*
- Assisted in leading multiple academic/social/networking events for many companies and hundreds of students
- Delegated features/improvements for a Discord bot with a currency, coffee chat, and anti-spam system for 2500+ members

Open Source Contributor | Microsoft Research Asia   **May 2021 - March 2022 | Beijing, China**
- *Python, HTML*
- Refined code scripts and translated modules involving fundamental AI knowledge from Chinese into English for ai-edu

Hackathon Lead Organizer + Mentor | SpartaHacks   **January 2019 - September 2019 | Guelph, ON**
- *Python, JavaScript, HTML, Docker, Blockchain*
- Planned event and attained major sponsorships, ensuring event success at no-cost for all 120+ participants
- Ran introductory Blockchain, Docker, and Python workshops to audiences with various computing skill levels