

Bill Jingwei Li

Github& Linkedin: cqlijingwei (+86) 13996456959-jingwei.li@mail.utoronto.ca

EDUCATION

University of Illinois at Urbana-Champaign, Illinois, IL	Master of CS	GPA: 3.9/4.0	Dec 2019
University of Toronto, Toronto, ON	Bachelor of CS	GPA: 4.0/4.0	Aug 2018

WORK EXPERIENCE

Senior Machine Learning Engineer, Boulder AI , Hangzhou, China	Sep 2021 – May 2024
<ul style="list-style-type: none">Constructed large language models (LLMs) chatbots to interact with customers, reducing 70% of labor communication cost.Fine-tuned LLM to improve 30% accuracy on customer responses and formed 20% of useful business rules and expertise.Developed OCR models using RNNs and NLP on paper recognition, achieving 90% accuracy and reducing 82% labor time.Used NLP word segmentation for company name parsing with 80% auto-labeling accuracy, reducing 90% labor time (C++).Built data analytics index system from scratch and Text-to-SQL model with Power BI, ETL, S3, Lambda and Redshift.Transformed MySQL databases between companies using data mapping and cosine similarity, improving 70% efficiency.Employed time-series models like ARIMA to forecast 1-3 months product sales with less than 5% bias using C#, .NET.Managed a 10-member team, deploying AI strategies(deep reinforcement learning) into production.Optimized operations research algorithms such as APS in Python, increasing production efficiency by 20%.Improved API response rate by 500% using Redis for caching asynchronous info.Combined diverse data sources (NoSQL, MySQL, byte files) into DocumentDB with 98%+ efficiency.	
Machine Learning Engineer, Alibaba Cloud , Hangzhou, China	July 2020 – Sep 2021
<ul style="list-style-type: none">Developed ensemble learning (CART, CNN) to optimize bus transit scheduling, reducing passenger wait time by 35% and operation time by 5%.Applied time-series model to predict parking availability with 90% accuracy in High Moral map services.Utilized statistical and time-series models to improve IoT anomaly detection from rule-based 30% to 85% accuracy.Implemented PCA and stat model correlation analysis using Python, Flink, PostgreSQL to improve generalized anomaly accuracy from 85% to 95%.Used Tensorflow and feature engineering to find high-value customers, increasing Cloud product sales revenue by 30%.Gained certifications: ACP Data Analyst, Cloud Computing, Big Data (match AWS certificates).Optimized ETL pipelines by 30%+ using AliCloud products DTS, MQ, Kafka, Blink and ODPS.Increased TB-level IoT data stream cleanness from 75% to 95%+ with Python, Java, PostgreSQL, OTS.	
Software Development Engineer, Cohesity , Mississauga, Canada	Sep 2019 – May 2020
<ul style="list-style-type: none">Built up operations automation pipeline for internal object creation, domain registration, software configuration, email automatic forward using Golang, bash shell script, HTML and VanillaGinkgo.Measured bridge performance by vdbench and fio testcases daily automation with analytics.Configured cluster nodes by force install with auto-generation of performance reports in Jenkins.Constructed data pipelines to integrate JSON, XML format files in MongoDB.	
Machine Learning Engineer Intern, Alibaba Cloud , Hangzhou, China	May 2019 – Aug 2019
Full-Stack Developer, Ministry of Education , North York, ON	May 2018 – Dec 2018
<ul style="list-style-type: none">CRUDED data from SQL databases. Implemented app screening functions in Spring and optimized OSAP admin service using Docker, K8s, Java, JavaScript, XML.Maintained financial payment server using Spring Boot, MyBatis with analysis of real-time noisy data.	
Data Engineer, IESO , Oakville, ON	May 2017 – Aug 2017

SKILLS

- Machine Learning**: Deep Learning, Reinforcement Learning, Ensemble Learning, NLP, Anomaly Detection, Time Series
- Data Science**: PyTorch, Scikit-Learn, TensorFlow, Pandas, Numpy, SciPy, NLP, Flink, Hive
- Programming**: Python & SQL (6 yrs), Java (5 yrs), R & SAS (4 yrs), JavaScript (3 yrs), HTML & CSS, MatLab, C++
- Cloud Computing**: AWS (S3, ECS, EMR, Lambda, Redshift), Azure, GCP, ACP(Alibaba Cloud Professional)
- Databases & DWH**: MySQL, NoSQL (MongoDB), OLAP (DocumentDB), Graph Database, Distributed File System