

GEOFFREY WONG HIN

Personal Website: <https://geoffreywonghin.vercel.app> | github.com/RyanStark223232

+1 825 945 7174 | geoffreywonghin@gmail.com | www.linkedin.com/in/geoffreywonghin | T2N 5A6, Calgary, AB

TECHNICAL SKILLSET

- Pytorch, Keras, TensorFlow, Scikit-learn, MLFlow, Tableau, PowerBI, RAG, LoRA
- Python, SQL, R, C, Java, GCP, Azure, Hadoop, Spark, Apache Beam, Databricks, React JS, Flask, FastAPI, Airflow, CICD

RELEVANT WORKING EXPERIENCE

RESEARCH ASSISTANT

THE CHINESE UNIVERSITY OF HONG KONG

JUN 2023- Aug 2024

- Optimized prediction and avoided overfitting in Deep Neural Networks (DNN) despite a data size of fewer than 300 by Implementing data augmentation such as Autoencoding and fast Fourier transform to video data.
- Developed chatbots with up to 95% accuracy in autonomously querying SQL databases and generating graphs based on accurate data utilizing an agent-based Large Language Model (LLM), capable of mirroring the abilities of a human data scientist based on the company's past working cases.

DATA SCIENTIST

HOTMOB LIMITED

JUL 2021- Jun 2023

- Boosted data extraction capacity by more than 10 times with half of the cost and extended website understanding capability from keyword matching to sentiment analysis, interest classification, and keyword extraction by leveraging transfer learning with GPT-2.
- Improved scalability and memory usage by 14x and reduced the processing time by 50% in our data pipeline by voluntarily identifying inefficiencies in the implementation of max connected component algorithm and devising solutions with heuristic algorithms.
- Removed up to 3 hours of manual labor daily by fully automated data processing pipelines using Airflow, allowing autonomous monitoring of data quality, updating ML models, and generating reports through dashboards.
- Brought in new revenue streams through a partnership with a global insurance company by impressing their Business Intelligence team in projects of upgrading their data pipeline to Azure, incorporating interpretable ML results, and improving cross-department ML model deployment.

EDUCATION

MASTER of DATA SCIENCE & ANALYTICS UNIVERSITY OF CALGARY

SEP 2024 - PRESENT

- Engineered End-to-end data pipeline capable of handling streaming data ingestion, app recommendations, and data dashboard on Google Cloud Platform (GCP)

BACHELOR of SCIENCE in COMPUTER SCIENCE THE CHINESE UNIVERSITY OF HONG KONG

SEP 2017 - Jul 2021

- Matched the accuracy and F-score of the 2017 PhysioNet Competition winner in detecting irregular heart rhythm with an Ensemble Convolutional Neural Network in Tensorflow.
- Discovered room for improvement in the best existing algorithms for the card guessing game by evaluating the randomness of card shuffling techniques using Recurrent Neural Networks in Keras.

PROFESSIONAL DEVELOPMENT & CERTIFICATES

Professional Machine Learning Engineer GOOGLE CLOUD CERTIFIED

JUN 2023 - Jun 2025

Professional Project Management Certificate GOOGLE

JUL 2023 - PRESENT