

# Kevin Wang

Phone: (647)-210-7930  
Email: [kevinwang7749@gmail.com](mailto:kevinwang7749@gmail.com)  
Website: <https://dungwoong.github.io/>  
LinkedIn: <https://www.linkedin.com/in/im-kevin-wang/>  
GitHub: <https://github.com/dungwoong>

## SKILLS

---

**Software Tools:** Python (Pytorch, pandas, Langchain, scikit-learn, plotly), SQL, Docker, AWS, R, Java, JavaScript/Node.js(React, NextJS, prisma), C/C#

**Data Science:** Regression (GLMs), Tree-based models (Random Forests, XGBoost), Computer Vision (CNNs), Language Models and Transformers, Cloud computing (Lambda Labs), Hypothesis testing

**Soft Skills:** Scrum/Agile processes, Jira, Git/Github, Linux

## EDUCATION

---

University of Toronto - Data Science Specialist April 2025  
❖ 3.95/4.0 GPA Toronto, ON

## EXPERIENCE

---

Data Science Research Intern May. 2023- Aug. 2024  
Pason Systems Corp. Calgary, AB

- ❖ Created and helped patent an algorithm for drilling dysfunction detection, increasing true positive predictions by 50% and reducing false positives by 40%. Prototyped using **Python** and productionized using **Java**. Created a **React** application to streamline the data labeling process.
- ❖ Maintained **KPI generation** program, designed new algorithms for cleaning and analysis of drilling data, and implemented in Java/MATLAB. Handled customer requests and facilitated cross-team solution design.
- ❖ Investigated machine learning applications in text data in the company. Explored the effectiveness of **Language Models** in interpreting and classifying drilling reports.

Junior Data Scientist/Analyst May. 2022 - Aug. 2022  
Neobi Technologies Calgary, AB

- ❖ Used **MySQL** to preprocess and perform **quality assurance** on thousands of cannabis data records for multiple clients. Praised for cleanliness and quality of the data.
- ❖ Developed web crawlers and data processing pipelines using **Selenium** and **Python**, used **PowerBI** to display analytics and help the sales team deliver clear and insightful visuals to clients

## PROJECTS

---

Car Listings Data Analysis - R

- ❖ Collected used car data, combining multiple data sources and performing **web scraping**
- ❖ Visualized data, and analyzed data by interpreting results of **Linear Regression**, and **XGBoost**
- ❖ Created a github website to display findings, adding reproducible descriptions of analysis

ReAct agent and RAG using LangChain - LangChain, OLLaMa

- ❖ Recreated the ReAct agent using Langchain, capable of browsing the web and executing calculations in Python
- ❖ Created a simple RAG system in Langchain to ask an LLM questions about an uploaded document
- ❖ Experimented with **prompt engineering** to fully leverage a smaller LLaMa3.2 model

Game AI using Imitation Learning with CNNs - PyTorch

- ❖ Used a **Convolutional Neural Network** to create an AI agent for the game Temple Run 2.
- ❖ Explored the **ShuffleNet** architecture, regularization techniques to boost in-game performance of the model.