

Kevin Wang

Phone: (647)-210-7930
Email: 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, R, Java, JavaScript/Node.js(React, NextJS, prisma), Docker, AWS, C, C#

Data Science: Language Models and Transformers, Tree-based models (Random Forests, XGBoost), Computer Vision (CNNs), GLMs, Naive Bayes, GMMs, PCA, PowerBI

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

EDUCATION

University of Toronto - Data Science Specialist - 3.95/4.0 GPA April 2025, 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 **deployed on 15+ rigs** using **Java**. Created a **React** application to facilitate labeling among 3 drilling engineers.
- ❖ Led ML drilling roadmap automation project, prototyped in Python, assisted in documentation and handoff process to two product managers
- ❖ Maintained **KPI generation** program, proposed algorithms for cleaning and analysis of drilling data, and implemented in Java/MATLAB. Handled customer requests and releases for **5+ account managers**.
- ❖ Investigated machine learning applications in text data in the company. Trained a **Language Model** from scratch in **PyTorch** on 150k+ drilling report entries.

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

- ❖ Used **MySQL** to preprocess and perform daily **quality assurance** on thousands of cannabis data records for 5+ clients. Maintained 50+ SQL scripts to be used for automatic cleaning.
- ❖ Developed web crawlers and data processing pipelines using **Selenium** and **Python**, used **PowerBI** to display analytics and worked with 2 clients to integrate requested visualizations.

PROJECTS

Car Listings Data Analysis - R

- ❖ Collected and visualized 900+ entries of used car data. Combined Kijiji data and **web scraped** metrics from MotorTrend.com. Analyzed effects of price decay using **Linear Regression**, and **XGBoost**
- ❖ Created a github website to display findings, adding reproducible descriptions of analysis

LLM Autonomous Web Browser - LangChain, Selenium, GPT4o

- ❖ Created an **LLM Agent**, capable of interacting with a Chrome browser by enhancing the WebVoyager paper.
- ❖ Replaced ReAct setup with LLMCompiler, and added a **multi-agent setup** for better validation.
- ❖ Experimented with prompt engineering, created evaluation tasks for the model.

Game AI using Imitation Learning with CNNs - PyTorch

- ❖ Used a **Convolutional Neural Network** to create an AI agent capable of achieving 100k+ score in Temple Run 2. Explored the **ShuffleNet** architecture, and multiple regularization techniques.