

Darwin Ye

+1 818-445-7824 | dxye@uwaterloo.ca | LinkedIn: darwin-ye-uwaterloo | GitHub: DarwinYip2022

EDUCATION

- **Northwestern University** | *Master of Science in Machine Learning and Data Science*
- **University of Waterloo** | *Bachelor of Mathematics in Honours Statistics & Honours Actuarial Science*

SKILLS

- Language: Python (Scikit-learn, TensorFlow, PyTorch), R, C, SQL
- Tool: AWS, Docker, Hive, Spark, Hadoop, Tableau, Git

PROFESSIONAL EXPERIENCE

Machine Learning Engineer | *Zebra Technologies* Sep. 2024 – Dec 2024 | Chicago, Illinois

- Developed deep learning models (**CNN** and **transformers**) to convert printer label images to ZPL code by training on a diverse, annotated dataset, resulting in accurate automated label generation.
- Designed and implemented an **NLP** interface using **GPT** models to allow users to modify ZPL code via natural language instructions, enabling dynamic, user-driven label customization.
- Led data preprocessing by standardizing images and augmenting datasets, improving model performance and ensuring reliable conversions across various label formats.

Data Science Consultant | *Nabors Industries* Sep. 2023 – May 2024 | Chicago, Illinois

- Utilized **LangChain** to develop an AI chatbot that enhanced data accessibility and user interaction by 30%, by processing natural language queries, retrieving relevant information, generating SQL code with a text-to-SQL **LLM**, and displaying results on a web platform.

Machine Learning Engineer | *Definity Financial* Sep. – Dec. 2022 | Toronto, Ontario

- Improved runtime efficiency for data cleaning and model-building by 40% through developing an R-based data **pipeline** using the {targets} package.
- Leveraged the {ROracle} package in RStudio Cloud to enhance data warehouse interactions and streamline data transfers, resulting in 30% faster data processing times.
- Standardized data preparation for predictive modeling from diverse sources, enhancing accuracy and consistency in model outputs with the {dplyr} package.

Data Scientist | *Intact Financial* Jan. – Apr. 2022 | Toronto, Ontario

- Boosted policy system efficiency by 25% by enhancing the commercial lines underwriting process through data consolidation and analysis using **SQL**.

Data Engineer | *Deloitte* Sep. – Dec. 2020 | Toronto, Ontario

- Built a data visualization web app **server** and **user interface** using the {Shiny} package in RStudio, providing users with advanced data visualization tools.

Data Scientist | *Canada Life* Jan. – Apr. 2020 | London, Ontario

- Created features for predictive modeling and developed interactive **Tableau** dashboards, leading to improved data insights and reporting capabilities by 50%.

PROJECTS

Drowsy Driver Detection System | *Python* Apr. – May 2024 | Chicago, Illinois

- Developed a deep learning-based drowsy driver detection system, utilizing **YOLO**, **CNN**, and **EfficientNet** models for real-time analysis, achieving high accuracy and precision in detecting drowsiness from video frames, thereby enhancing road safety through proactive monitoring.

Cloud Engineering | *Python* Apr. – May 2024 | Chicago, Illinois

- Designed and implemented an **AWS** data processing pipeline utilizing **S3** for storage, **ECR** for Docker container registry, **ECS** for container management, and **EC2** for infrastructure, automating deployment of Dockerized models and delivering real-time analytics through **Streamlit**, significantly enhancing data processing efficiency and scalability.

Recommendation Engine | *Python* Jan. – Mar. 2024 | Chicago, Illinois

- Built a **Python**-based recommendation engine for e-commerce, improving user engagement by 20% and sales conversions by 15% through **collaborative** and **content-based filtering**.