

Roman Wu

✉ roman.y.wu@gmail.com | ☎ 289-834-0229 | github.com/roman-y-wu | [linkedin.com/in/roman-y-wu](https://www.linkedin.com/in/roman-y-wu)

EDUCATION

Northern Arizona University

Master of Computer Science (Remote) (GPA 3.9/4.0),

Flagstaff, AZ

2025 – Present

University of Waterloo

Bachelor of Mathematics President's Scholarship (GPA 3.83/4.0)

Waterloo, ON

2020 – 2023

SUMMARY

- * A motivated AI/ML engineer with 3 years of experience delivering scalable ML and software products across industries.
- * MLOps in Consulting (Sojo Canada), Data Scientist in InsurTech (Co-evolves), Machine Learning in Finance.
- * Experienced in software development, SaaS, analytical skills, and a solid grasp of ML algorithms, data visualization, data structures, OOP, cloud computing, version control, operations research, REST APIs, data engineering and LLMs.

SKILLS

Programming Languages: Python, SQL, R, Coding, C++, Java, JavaScript, MATLAB, SAS

Frameworks: Pandas, NumPy, Scikit-learn, TensorFlow, PyTorch, Spark, Databricks, Snowflake, Azure, Docker, Kubernetes

Technologies: Statistics, AI/ML, A/B Testing, KPIs, Business intelligence, Advanced analytics, MLOps, Data manipulation, Sales

EXPERIENCES

SOJO Canada

Toronto, ON

Data Scientist

Sep 2024 – Present

- Optimized the existing PII detection pipeline using Bert and graph neural nets, improving compliance accuracy.
- Refactored GAM/XGBoost algorithms and CV validation to estimators using OOP techniques.
- Optimized KMeans + XGBoost models for customer segmentation, increasing marketing campaign conversion by 12%.
- Optimized and monitored auto insurance ML pricing pipelines for Ontario and BC, deployed the models with ONNX runtime and H2O, enabling the pricing team to use production ML rating algorithm.
- Optimized feature stores to unify ML pipelines across regions to support underwriting decisions, compatible with Snowflake.

Co-evolves

Toronto, ON

Data Scientist

Dec 2023 – Aug 2024

- Conducted inferential tests and experiments design to analyze underwriting risks on renewal probability.
- Built a fully automated OOT model monitoring dashboard using airflow and dbt, enabling batch-driven data refresh, visual diagnostics (Gini, lift, one-way plots), and accelerating model degradation detection and retraining decisions by 40%.
- Designed a **RAG** NLP system and conversational chatbot supporting the call center and allowing agents to query using natural language.
- Overhauled the API ingestion pipeline using vendor's REST API endpoints, processing over 100k call recording data per day to mitigate company's loss of functionality of compliance reviews while maintaining vendor relations.
- Created profit variance and cost-control dashboards to track expected vs. actual margins, and loss cost, providing weekly insights that guided pricing adjustments and operational decisions.

Harvest Fund Management

Beijing, China

Machine Learning Engineer Intern

Sep 2022 – Dec 2022

- Developed hybrid ML models (Bayesian Regression, CNN, LSTM) to forecast claim costs and quote conversions using Keras/PyTorch, reducing pricing errors by 12% and increasing high-risk portfolios identification by 15%.
- Optimized PySpark ETL pipelines by integrating AWS Step Function, Lambda, and Polars, boosting processing speed by 20%.
- Developed Python BI modules for risk scoring and Tableau dashboards, reducing manual analysis by 30% while optimizing MySQL queries with indexing.

PROJECTS

Kaggle Competition (Gold Medal – E-Commerce Classification)

Sep 2022 – Dec 2022

Performed classification on both texts and images of e-commerce datasets by training transformer and MLP classifiers, combined with supervised models, such as KNN, Random Forest using Scikit-learn.

Neural Machine Translation using Transformer

Sep 2022 – Dec 2022

Built Transformers from scratch, implementing core components like multi-head attention, beam search, and encoder-decoder architecture for effective multi-lingual translation and retrieval tasks.