Andy Yang

+1-647-461-6466 • syyyang003@gmail.com • [Personal Website](#) • [LinkedIn](#) • [GitHub](#)

SUMMARY

- Over 3 years of Experience in Data Analytics at Companies such as TD, and Toyota with projects in implementing ETL pipelines, Python and R full stack applications, and innovative machine learning validation solutions
- Completed machine learning projects involving regression and classification modelling, neural network optimization, R and Python dashboard development, and model deployment in AWS and Spark
- Strong coding skills in Python (Spark), R, SQL, with proficiency in dealing with large datasets using Hadoop

WORK EXPERIENCE

TD Bank Group *Toronto, ON*

Senior Quantitative Analyst (Model Validation)   Apr 2023 - Present

- Engineered an innovative Python application that automated the consolidation of model validation KPIs into Tableau dashboards using a streamlined SQL data pipeline, expediting the quarterly review process by over 2 weeks
- Elevated data science workflows by leading the development of Python and PySpark model validation code to supersede existing SAS repositories and seamlessly integrate with PySpark and Hadoop cloud environments
- Collaborated with model validators to script technical improvements to R Shiny dashboards, improving data visualization clarity and implementing predictive modelling analytics such as stress and sensitivity testing

Quantitative Analyst (Model Validation)   Sept 2022 – Apr 2023

- Validated a variety of machine learning and AI fraud detection models with SAS and Python, conducting comprehensive performance testing including AUROC, K-S, hyperparameter optimization, and stability analysis
- Showcased critical thinking ability by conducting thorough assessments of ML model integrity, encompassing model methodology, feature engineering, explainability and reproducibility

Royal Distributing Inc. *Guelph, ON*

Implementation Specialist   Apr 2021 – Sept 2021

- Orchestrated the integration of custom API and web data from 10+ retail vendors into the company ETL pipeline using C# and .NET framework

Toyota Motor Manufacturing Inc. *Cambridge, ON*

Information Systems (Co-op)   Jan 2020 – Apr 2020

- Displayed excellent analytical ability by optimizing SQL server database schema and restructuring ETL data pipelines, reducing server load and speeding up query response times by up to 300%

PROJECTS

LiDAR Object Detection and Classification for Cities   May 2022 – Jun 2022

Reliance Foundry Co. Ltd. UBC MDS Capstone

- Trained and optimized convolutional neural network using PyTorch to perform 3D object detection within LiDAR point clouds, achieving a maximum mean average precision score of 0.96
- Recognized as one of the top 3 Capstone projects of the UBC MDS 2023 Cohort

EDUCATION

M.S. of Data Science   Sep 2021 – Jul 2022

University of British Columbia, Vancouver, BC ([coursework information](#))

BASc. of Chemical Engineering   Sep 2016 – Apr 2021

University of Waterloo, Waterloo, ON

SKILLS

Coding Languages and Tools: Python, R, C#, Jupyter Lab, PySpark, HTML, CSS
Statistics: Hypothesis Testing, Segmentation Analysis, A/B testing, Full Factorial Design, Generalized Linear Modelling
Data Platforms: SQL, NoSQL, PostgreSQL, Hadoop, Tableau, Power BI, AWS, Spark, SAS (certified in 9.4)
Machine Learning: Scikit-Learn, PyTorch, Ensemble Modelling, Neural Networks, Natural Language Processing