

Shirley(Can) Cui

Shirley.cui123@gmail.com | Canada | <https://github.com/cc121-shirley>

SUMMARY OF QUALIFICATION

- U of T HBS Major in Statistics and Economics, 5 years + experience in Data Science area
- 3 years data science and model development experience in banking credit risk area
- Data Science Skills: Exploratory Data Analysis, Data Cleaning, Data Preprocessing, Data ETL, Feature Engineering, Machine Learning & Predictive Model (Regression, Random Forest, XGBoost) , Data Visualization, Statistical Analysis (Probability, Hypothesis Test, Regression Analysis), A/B Testing
- Technical Skills: Advanced level of **Python (Pandas, Numpy, Scikit-Learn, Scipy, H2O, Matplotlib etc)**, Pyspark, SQL, **Power BI / Tableau, R, GIT**

EDUCATION

Honor Bachelor of Science in Statistics and Economics, **Sept 2015-Jun 2020**
University of Toronto

- Relevant courses:
Introduction to programming (Python), Regression analysis, Machine learning, Surveys,sampling and observational data, Design and analysis of experiments,

WORK EXPERIENCE

Applied Machine Learning Scientist **Nov 2022-Now**
TD Bank

- Handle large databases of credit risk related historical customer data and perform data ETL for modelling using SAS,SQL
- Leverage advanced data analytics to derive insights and develop credit risk scorecard model using Machine Learning and statistical methods (Feature Engineering, Hyperparameter Tuning, Cross Validation, XGBoost, Model Evaluation,Shap Ranking) in cloud environment (Azure Databricks) to assist business partners achieving growth
- Refine machine learning model pipeline and achieve 20% performance improvement
- Lead credit risk model migration for 5+ models from SAS to Azure and reduce the running time by 5x by optimizing codes
- Write model development reports on end to end model development process for stakeholders to review
- Work closely with model validation & monitoring team to support all stages of model validation on scorecard models, ensuring model performance stability
- Cross-Functional Collaboration: provide decision support, analytical, and statistical service through modeling and analysis to help strategy teams to achieve business objectives in credit risk area

Senior Analyst , Pharmacy Planning & Insights **Sept 2021-Nov2022**
Shoppers Drug Mart

- Lead data analytics projects and generate insights to solve business problems and drive the strategy and decision making
- Extract, synthesize large amount of pharmacy data from multiple sources (SQL, Microstrategy) including transaction, customer and services data
- Use Python to apply data science & statistical analysis techniques for projects and automate executive reporting
- Develop dashboards from Power BI for business insights and enhancing the reporting system by automating data extraction and manipulation
- Build professional decks on multiple projects and delivering the presentations to senior management team

Machine Learning Associate **Nov 2020-Nov 2021**
University Health Network (KITE Research)

- Work on CLSA research project using Machine Learning & Data Analytics techniques on Python to handle large dataset independently
- Apply Exploratory Data Analysis, Data Cleaning, Data Visualization using Python packages (Pandas, Numpy, Matplotlib)
- Handle missing values, mixed data (Mean, Mode, EM, MICE Imputation, Feature Hashing) by different techniques
- Perform Statistical Test, Feature Selection and Machine Learning Algorithm (Logistic Regression, Random Forest, XGBoost, Cross Validation) on Scikit-Learn, H2O,Scipy
- Develop charts (Matplotlib, Seaborn), reports, slides and report to Machine Learning scientist and clinical scientist