Yumi Si

Phone: 3433630369   Email: yumeng.s@queensu.ca
Address: 38 Greville, Toronto   Personal website: https://yumengqueensu.github.io/

TECHNICAL SKILLS

Programming & Modelling: Python, Java, C#, R, SAS, SQL, pandas, NumPy, Linear Regression, Logistic Regression, XGBoost, feature engineering & selection, cross-validation, basic PyTorch

Financial Engineering: Monte Carlo (option pricing incl, scenario ECL); VaR/ES with Kupiec & Christoffersen tests; IFRS 9 (PD/LGD/EAD), WOE/IV scorecards,

Tools & Visualization: Streamlit dashboards; Plotly/Matplotlib; Git/GitHub; Docker; Linux; PostgreSQL/MySQL

INTERNSHIP EXPERIENCES

China Electronics Corporation Co.,   Oct 2020 – Feb 2021
Model Developer Intern
- Contributed to an SME loan-application risk model, implementing the KMV (distance-to-default) framework for probability-of-default estimation and ongoing model monitoring/calibration
- Built the training data pipeline from bank feeds: data cleansing and outlier handling, standardisation, and feature engineering/selection
- Responsible for the interaction between the front-end and the database to enable model deployment

Baidu Cloud Computing Technology (Beijing) Co.,   Mar 2022 – Jun 2022
Data Systems Algorithm Engineer Intern
- Designed and implemented core data-processing algorithms in Python: binary search tree for fast indexing/lookup, sliding-window time-series transforms, a modular feature-extraction pipeline
- Improved search efficiency by replacing linear scans with BST queries ($O(n) \rightarrow O(\log n)$).
- Contributed to UI and visualisation, making model outputs clear and accessible for non-technical users

PROJECT EXPERIENCES

IFRS 9 ECL Modeling and Stress Testing Framework for Canadian Retail Credit
- Built an IFRS 9 compliant ECL system for Canadian retail loans using a SQL – SAS – Python stack. Trained and validated on 10,000 records from Statistics Canada/Bank of Canada to compute forward-looking expected losses.
- Processed payment histories, extracted 30+ risk features, developed PD/LGD models with logistic/Beta regression, and implemented three-stage credit deterioration tracking for provisioning
- Integrated macro stress testing to enable dynamic risk adjustments during market downturns.

Fraud Strategy A/B Backtesting and Decision Optimisation
- Built a Python-based card-fraud strategy A/B backtesting platform with Logistic Regression and YAML-based strategies to compare model+rule combinations on consistent time slices and choose cost-optimal actions.
- Evaluated using AUCPR and Recall at fixed false-positive rates.
- Delivered a three-way decision flow (block/review/allow), governance artefacts and a Streamlit dashboard, reducing expected losses per 1,000 transactions in scenario tests.

Quantitative ETF Allocation: Markowitz – to – Deep-Learning Optimisation and Evaluation
- Designed a systematic portfolio allocation project benchmarking classical and ML strategies. Built in Python on 5 ETFs (SPY, IWM, EEM, GLD, TLT) with ~1,760 daily observations data for allocation and testing
- Compared traditional Markowitz method with Lasso, Decision Trees, XGBoost, etc.; the Gated Recurrent Unit model achieved a ~0.1 higher Sharpe Ratio than classical Markowitz
- Delivered interactive visualisation with Streamlit/Plotly

EDUCATION

Queen's University, Kingston, Canada   **Sep 2022 – Jun 2024**
Master of Engineering in Electrical and Computer Engineering | 4.3/4.3
Tianjin Normal University, Tianjing, China   **Jul 2016 – Jun 2020**
Bachelor of Engineering in Software Engineering
Key courses: Python/Java/C# Development, SQL Database, Machine Learning, Statistics, Probability Theory