Machine Learning for Model Development in Market Risk

P Phani Kumar

Mar 2018

Introduction

Machine learning as part of Artificial Intelligence, famously known as AI has become a buzz word across the industries in recent times. Machine learning has been prevalent for a long time in academia and research though it has occupied prominence in the recent times. Banking and Financial services industry has also witnessed the need for understanding machine learning and implementing it in various activities viz. AML/KYC, Reporting, Modeling activities etc.

This paper details the scope of machine learning in model development along with challenges that are involved in the process of implementation of the framework, with focus on market risk domain. It initially outlines the characteristics of modeling/reporting in market risk including trading and banking book portfolios and then gives an overview of the model development process. Finally, it provides the possible areas where machine learning can be employed in the market risk.

Market Risk Model Development

The models developed for measuring market risk of a bank can be classified in various ways and one of the classification is based on purpose i.e. cash flow forecasting models viz. exposure forecasting, prepayment, runoff models etc. and discount rate estimation models (Interest rate modeling). Another way of classification of the models is based on the book viz. models for banking book and models for trading book though some of the models developed are used for both the banking and trading book (e.g. yield curve construction).

The models for banking book can be of various types and include models to estimate runoff and pricing of liabilities viz. time deposits and non-maturity deposits, prepayment of retail/commercial assets of the bank etc. These are typically behavioral models and focus on capturing the behavioral trends in the portfolios. These behavioral models are regression based models and there is no limitation on the independent variables that can be used to estimate the dependent variables viz. developer needs to capture the various variables that can impact the behavior in the portfolio and they may be unique for each of the portfolios. However, it is to be noted here that approach in identifying the variables may be similar across portfolios.

Models for trading book primarily can be either for pricing of plain vanilla or structured trading securities viz. swaps, forwards, options etc. or for VaR estimation of these products. These models primarily involve understanding the underlying stochastic processes and then model the required variable. This paper does not delve in to further details on these type of models as the focus of this paper is on illustrating the scope of machine learning for development of these models.

In this regard, model development process in a bank can be broadly divided in to five stages.

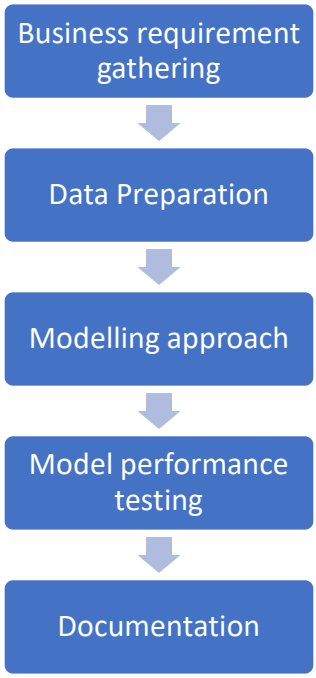

Figure 1 Model development Process for a financial institution/bank

Stage 1: Understand objective/ purpose of the model and scope of the model which is generally described in BRD i.e. business requirement document. It includes the portfolios and the desired output of the model for business user
Stage 2: Dataset preparation which is the most critical and time taking aspect of any of the model development process.
Stage 3: Deciding on modeling approach/methodology based on underlying data and desired output and apply statistical/stochastic theory for developing the model
Stage 4: Back testing the model output with out of sample data and finalize the model along with overlays based on model performance in out of sample period
Stage 5: Documentation of the entire model development process

Employing any of the machine learning techniques for the entire process currently is highly challenging due to inherent expertise required while constructing the models and from regulatory perspective. Further, validation of these models is challenging and the current framework for validation needs to be revisited. However, implementing them partially in one of the stages of process is much easier and is being tried at different levels by various banks.

Application

Stage 1: Business Requirements

Scope of gathering the requirements and understanding objective of the model through the artific ia l intelligence techniques is far-fetched and there is a lot of headway before implementing AI in this stage.

Stage 2: Dataset preparation

This stage appears to have huge scope for implementation of machine learning techniques. It is beyond doubt that a lot of automation is being performed in this phase of model development by various banks. However, the implementation of machine learning techniques is in nascent stage and is being explored in recent past. This stage involves data collection, cleansing and preparing the training and testing data set. In the case of banking book models, the data collection phase will involve collecting data for all the relevant variables including macro-economic variables and accordingly the machine learning techniques can be employed to collect and cleanse the relevant macro variables data based on the estimator that needs to be estimated by the model. In case of trading models, this stage involves identifying the relevant instruments trading in the market and sourcing the relevant market data for them.

Stage 3: Modeling Approach

This stage involves designing the model approach that can be employed to develop the model. With respect to banking book models, employing the linear/logistic regression modeling techniques as part of modelling approach is prevalent for long periods of time and focus is on employing the machine learning techniques for improving the forecasting accuracy. It has been observed that machine learning techniques are readily being employed for improving accuracy in PD models in the credit risk viz. decision trees, Random Forest and boosting based models. Accordingly, there is a huge scope for employing them in market risk banking book models. In the case of trading models, machine learning is already being used for developing trading strategies. Accordingly, trading models in the regulatory space can employ the machine learning practices. However, one of the major issues that is being observed while employing machine learning techniques is that it tends to overfit the data which needs to be addressed.

Stage 4: Model Performance Testing

Out of sample back testing is performed to estimate the model performance with testing dataset and typically the setting of benchmarks/thresholds for monitoring the performance is the salient point in this phase. There is a scope for implementing the machine learning techniques to set the same and perform the back testing along with performance monitoring.

Stage 5: Documentation

Documentation is last step in the model development and machine learning techniques are currently being employed for documentation of the model development project by employing NLG. For e.g. filling the relevant information by analyzing the graphs, definitions of various parameters, variables used in the model.

It can be observed from the above that there is a huge scope for implementing the machine learning in market risk. However, as these models are subject to validation, there is a need for cohesion across the organization i.e. development/validation and regulations for effective modeling practice using machine learning.

Author

P Phani Kumar is a consultant and a subject matter expert in the risk and regulatory domain at Genpact LLC. He has 7+ years of experience in the areas of risk management consulting and project management, working on consulting engagements with global financial institutions, advising clients on financial risk management.

He has been involved in multiple risk consulting engagements and implementation projects on both regulatory and strategic fronts in Basel II/III, CCAR, ALM areas for financial institutions in the US. He is also responsible for developing solution offerings and frameworks on Basel III, CCAR/DFAST for Genpact clients in the US banking sector.