Global Journal of Computer Science and Technology: D
Neural & Artificial Intelligence
Volume 18 Issue 1 Version 1.0 Year 2018
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Has Machine Learning Arrived for Banking Risk Managers?

By Manoj Reddy

Abstract- Machine learning is one of the most exciting and powerful cognitive levers out there in the Industry and today Risk managers are grappling to make sense of whether it is just a hype or does it really have a value to add in Banking Risk Management. The article attempts to give a brief introduction into foundational concepts of machine learning and highlights some of the problems with the current predictive models and also some of the most popular pilot or candidate Use cases for Machine learning adoption. It also highlights the critical success factor which one needs to consider or be aware of in in adoption of Machine leaning to solve business problems. The paper intends to demystify the conundrum called as Machine learning and elucidate it to an extent which will enable the new age Risk & Regulatory managers to carefully evaluate and decide on an adoption of Machine learning techniques and solutions to solve specific Business problems.

GJCST-D Classification: I.2 I.5 .6, I.2.m

HASMACHINELEARNINGARRIVEDFORBANKINGRISKMANAGERS?

Strictly as per the compliance and regulations of:

Has Machine Learning Arrived for Banking Risk Managers?

Manoj Reddy

Abstract- Machine learning is one of the most exciting and powerful cognitive levers out there in the Industry and today Risk managers are grappling to make sense of whether it is just a hype or does it really have a value to add in Banking Risk Management. The article attempts to give a brief introduction into foundational concepts of machine learning and highlights some of the problems with the current predictive models and also some of the most popular pilot or candidate Use cases for Machine learning adoption. It also highlights the critical success factor which one needs to consider or be aware of in in adoption of Machine leaning to solve business problems. The paper intends to demystify the conundrum called as Machine learning and elucidate it to an extent which will enable the new age Risk & Regulatory managers to carefully evaluate and decide on an adoption of Machine learning techniques and solutions to solve specific Business problems.

I. Background and Business Context

The Digital revolution is in the process of transforming the Banking Industry like never before. From understanding the customer proactively and deciding your next best action to gaining greater efficiencies and scale in running the Business, technology has played a huge role in enabling delivery of maximum business value to the customer in the shortest possible time. Artificial Intelligence is challenging the very realms of possibilities in how you run your Business and what seemed fiction or theoretical a few years back is now becoming a reality which Banks will need to consider seriously to stay in the game.

Machine Learning is powering the Artificial Intelligence revolution through its effort to make the predictive capabilities of Business lot more accurate, repeatable and scalable. Though the early adopters of Artificial Intelligence have been the customer centric functions of Bank, there has been a gradual progression towards exploring its potential to solve Banking Risk Management problems.

II. So, Are There Any Limitations with How Things Work Now?

One of the foundational principles of Risk management is the need to not only assess or quantify your risk, but also to do it early enough to effectively mitigate it. The Industry has relied very heavily on algorithm based and conventional statistical models driven by knowledge, experience and judgment of the Business Subject matter experts across banks globally. Though the Conventional Risk models have stood the test of time and have helped Banks identify and manage their risk, they are yet to achieve highest levels of accuracy and additionally they are not equipped to handle huge chunks of structured and unstructured data which potentially hold critical risk insights.

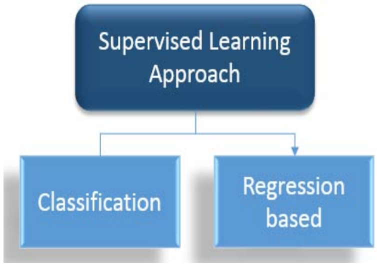
For Instance, The Industry average of the accuracy of a PD model or a Loan Default prediction model typically could be anywhere between 80 to 90%, and in most cases we are unable to scale up the model to process greater volumes of historical data due to technology and infrastructure limitations. Additionally since the variable selection and its corresponding Significance in prediction is largely humanly defined with some statically inferences, we at most times are not able to fully derive the benefit of drawing insights from the breadth and depth of historical data. Similarly Industry is still battling with high number of false positives in the Anti-Money Laundering space and is excessively erring on the side of safety which is hurting operating margins and efficiencies. There are number of such other predictive outcomes in Risk management where leveraging of additional insights could very positively increase the accuracies.

III. What is Machine Learning in the Context of Risk Management?

At a very high level there are two most popular approaches to machine learning they are namely supervised and Unsupervised learning.

Supervised Learning: This is an approach to Machine learning where the historical Inputs data is tagged with its corresponding Business outcomes and the Machine learning solution is expected to identify and learn the patterns in the inputs data associated with a Business outcome and self-develop an algorithm based on this learning to predict a Business outcome for a future instance. There are two primarily utilities of Supervised learning:

Author α: *Head of BFS Risk & Compliance, United States.*
e-mail: manoj.reddy@tcs.com

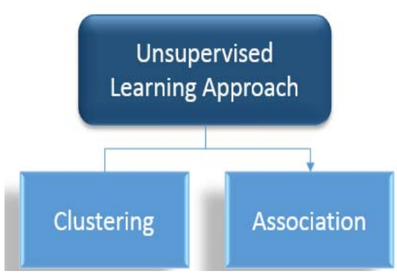


Classification: This is used for Business problems where the prediction deals with classification. For Eg: Whether a Loan will default or not. (Yes or No).

Regression: This is used for Business problems where a quantified value is being predicted. For example what is the Probability of Loan default? (Expressed in percentage).

In case of Supervised Learning it is important to have a good mix of historical data supporting both the positive and negative hypothesis both for the training set and a test set. It is recommended that we have them almost equal. Additionally it is recommended to have 80% of your historical data as Training and 20% as the Test set.

Unsupervised Learning: This is an approach to Machine learning where the historical Inputs data is fed into the machine learning solution without any tagging of the Business outcomes and the solution is expected to decipher or self-develop an algorithm for prediction based on its own interpretations of the patterns in the data without any guidance or indicators. The Primary difference between supervised and unsupervised learning is that the historical Input training data is tagged with Business outcomes as well in cased of supervised learning. There are two Primary utilities of Supervised learning:



Clustering: This is used for Business problems where records with homogenous characteristics are clustered together. For Eg: Customer Segmentation for Consumer Credit Risk Calculations.

Association: This is used for Business problems where the prediction is looking at the association of variables.

For Eg: Impact of increased draw-down on credit lines prior to default.

Depending on the problem we are trying to solve any of the above mentioned techniques can be used. The end objective is to leverage the ability of machines to crunch huge amounts of historical data and uncover the hidden insights which are beyond human compression and conventional statistical analysis.

IV. So What are Some of the Problems Machine Learning can Solve?

There is definitely a lot of buzz and excitement around leveraging Machine learning for predictive models needed by Risk Managers, but a clear understanding of the business value and the specific need for leveraging it should be carefully understood and evaluated before zeroing in on the USE cases. Leveraging machine learning for processes which are already fairly accurate with a Business rule or statistical models may not be good candidates for early adoption.

Following are some of the top candidate processes or USE cases being explored in the Industry for Machine Learning Adoption in the Banking Risk management space.

- *Anti-Money Laundering (AML) Transaction Monitoring:* Historical data on Suspicious transactions can be analyzed through machine learning techniques to understand patterns that are associated with suspicious transactions from those that are not.
- *Risk Based Credit Approvals:* For high dollar value credit approvals which are largely judgement based and not business rule driven, there could be certain factors or combination of factors which the underwriters might have used to approve them. Machine leanring could help analyze and interpret a pattern associated with approvasls and develop an algorithm to predict it more consistently.
- *Loan Default Prediction:* Some initial studies have shown that Machine learning techniques have been able to uncover some hidden insights and pattern of credit deterioration from historical data which can now enable us to identify early signs of credit deterioration or eventual default based on time series data of defaults. This specially has significance both from a strategic Loan Default management perspective and also from perspective of Regulatory expectations in IFRS 9 and CECL to identify early warning signals of Credit deterioration.
- *Risk Forecasting Models:* Similar to classification related problem statements, machine learning techniques can be effectively used for Regression based forecasting as well. Primarily, forecasting models for Probability of Default (PD), Loss Given Default (LGD) and Credit Conversion factor (CCF) can show greater levels of accuracies in forecasting the quantum of risk with greater degree of precision and Accuracy.
- *Consumer Loan Risk Segmentation:* Consumer loans need to be segmented in pool of loans exhibiting homogenous characteristics. This is

done based on rule based categorization today based on business knowledge and experience built over the years. However, clustering techniques in machine learning provide tremendous scope and potential for the consumer loan segmentation using methods such as KNN.

V. Critical Success Factors in Machine Learning Adoption

- It is important to understand that *different problems statement could warrant different machine learning techniques.* So there is no one size fits all. It is an iterative processes. Sometimes you might use a hybrid approach of using more than one technique as well.
- It is important to use actual historical data which carries the *real insights, and any attempts to synthetically create data may not be able to capture the idiosyncratic insights and pattern* for a given business outcome which your historical data might carry.
- A machine learning algorithm needs to be given time to improve its accuracy. The *more the historical data the deeper is the learning and greater will be the accuracy.*
- It should be viewed as *a Solution that is there to assist the Business and Subject matter experts and not to replace them.* They should at best recommend and not be given a final Business decision making role.
- There needs to be *gradual adoption of Machine learning solutions* as against a Big bang approach.
- *Machine learning solutions can be combined with Robotic Process Automation* (RPA) solutions to develop Cognitive RPA solutions which have tremendous business value and efficiencies.
- There needs to be at *the very least 4 quarters of parallel run* of the Machine learning solution with existing predictive models prior to moving them to production.
- *It is not purely a statistical or a technical exercise. Machine learning requires equal amounts of domain or subject focus* to identify the right data sources, data sets and data variables. The qualification of relevant data will optimize your iteration cycle.

VI. Conclusion

Machine learning is here to stay and is making rapid strides across multiple Industries. Though Machine learning as a concept has existed for more than three decades now, it's the advancement in Data storage and data management powered by the Big Data and Digital revolution has made Data mining more affordable and easier. There are number of Banks and financial institutions which have commenced their exploratory journey into machine learning world and for all new age Risk Mangers Machine learning is no more a work of fiction but is a reality with tremendous potential to proactively and effectively predict and manage their Risk. The Journey has just begun.