Article

Machine Learning in Banking Risk Management: A Literature Review

Martin Leo * , Suneel Sharma and K. Maddulety

SP Jain School of Global Management, Sydney 2127, Australia; sunil2500@gmail.com (S.S.); k.maddulety@spjain.org (K.M.)
* Correspondence: amaleo@yahoo.com; Tel.: +65-9028-9209

check for updates

Abstract: There is an increasing influence of machine learning in business applications, with many solutions already implemented and many more being explored. Since the global financial crisis, risk management in banks has gained more prominence, and there has been a constant focus around how risks are being detected, measured, reported and managed. Considerable research in academia and industry has focused on the developments in banking and risk management and the current and emerging challenges. This paper, through a review of the available literature seeks to analyse and evaluate machine-learning techniques that have been researched in the context of banking risk management, and to identify areas or problems in risk management that have been inadequately explored and are potential areas for further research. The review has shown that the application of machine learning in the management of banking risks such as credit risk, market risk, operational risk and liquidity risk has been explored; however, it doesn't appear commensurate with the current industry level of focus on both risk management and machine learning. A large number of areas remain in bank risk management that could significantly benefit from the study of how machine learning can be applied to address specific problems.

Keywords: risk management; bank; machine learning; credit scoring; fraud

1. Introduction

Since the global financial crisis, risk management in banks has gained more prominence, and there has been a constant focus on how risks are being detected, measured, reported and managed. Considerable research (Van Liebergen 2017; Deloitte University Press 2017; Helbekkmo et al. 2013; MetricStream 2018; Oliver Wyman 2017), both in academia and industry, has focused on the developments in banking and risk management and the current and emerging challenges. In tandem, there has been a growing influence of machine learning in business applications, with many solutions already implemented and many more being explored.

McKinsey & Co highlighted that risk functions in banks, by 2025, would need to be fundamentally different from what they are today. The broadening and deepening of regulations, evolving customer expectations and the evolution of risk types are expected to drive the change within risk management. New products, services and risk management techniques are being enabled through the application of evolving technologies and advanced analytics. Machine learning, identified as one of the technologies with important implications for risk management, can enable the building of more accurate risk models by identifying complex, nonlinear patterns within large datasets. The predictive power of these models can grow with every bit of information added, thus enhancing predictive power over time. It is expected that machine learning will be applied across multiple areas within a bank's risk organisation. Machine learning has also been recommended as an initiative that could help in the transformation of the risk management function at banks.

The paper seeks to study the extent to which machine learning, which has been highlighted as an emergent business enabler, has been researched in the context of risk management within the banking industry and, subsequently, to identify potential areas for further research. The aim of this review paper is to assess, analyse and evaluate machine-learning techniques that have been applied to banking risk management, and to identify areas or problems in risk management that have been inadequately explored and make suggestions for further research.

To determine the risks specific to banks, as an alternate to leveraging on existing literature, this paper provides a taxonomy of risks that is developed based on a review of bank annual reports. An analysis of the available literature was carried out to evaluate the areas of banking risk management where machine-learning techniques have been researched. The research evaluated the risk areas where machine learning has been implemented in the risk types and the specific risk methodology they addressed. The analysis also identified the machine learning algorithms being used, both for specific areas and in general.

Section 2.1 provides an overview of risk management at banks, the key risk types and risk management tools and methodologies. Section 2.2 gives a quick introduction to machine learning and its use. Section 3 begins by providing an overview of the research methodology. The section further examines the existing research around the application of machine learning in the management of risk at banks. It provides an analysis of the areas where the application of machine learning has been studied, highlighting areas where there is little to no academic study. Section 4 discusses the key observations from the review, stressing the potential challenges and topics that could be addressed in the future. Section 5 summarises the general findings from the study. The paper concludes by listing additional areas or problems in banking risk management where the application of machine learning can be further researched.

2. Theoretical Background

2.1. Risk Management at Banks

The bank's management's pursuit to increase returns for its owners comes at the cost of increased risk. Banks are faced with various risks—interest rate risk, market risk, credit risk, off-balance-sheet risk, technology and operational risk, foreign exchange risk, country or sovereign risk, liquidity risk, liquidity risk and insolvency risk. Effective management of these risks is key to a bank's performance. Also, given these risks and the role that banks play in financial systems, they are subject to regulatory attention (Saunders et al. 2006). The regulators require banks to hold capital for the many risks that arise and are carried due to a bank's varied operations. The Basel standards for the determination of capital requirements were developed in 1998, and since then, have developed and evolved. Capital is required for each of the main risk types. Credit risk has traditionally been the greatest risk facing banks, and usually the one requiring the most capital. Market risk arises primarily from the trading operations of a bank, while operational risk is the risk of losses from internal system failures or external events. In addition to calculating regulatory capital, most large banks also calculate economic capital, which is based on a bank's models rather than on prescriptions from regulators (Hull 2012). The main risks that banks face are credit, market, and operational risks, with other types of risk including liquidity, business, and reputational risk. Banks are actively engaged in risk management to monitor, manage and measure these risks (Apostolik et al. 2009).

Market risk can be defined as the risk of losses "owing to movements in the level or volatility of market prices" (Jorion 2007). Market risk includes interest rate risk, equity risk, foreign exchange risk and commodity risk. Interest risk can be defined as the potential loss due to movements in interest rates. Equity risk can be defined as the potential loss consequent to an adverse change in the price of a stock. Foreign exchange risk can be defined as the risk that the value of the assets or liabilities of a bank changes due to fluctuations in the currency exchange rate. Commodity risk can be defined as the potential loss due to an adverse change in the price of commodities held. The market risk framework

of the Basel accord consists of an internal models approach and a standardised approach. To capture tail risk better, the revised framework also saw a shift in the measure of risk under stress from the Value-at-Risk (VaR) to Expected Shortfall (ES) (Basel Committee on Banking Supervision 2006).

Credit can be defined as the risk of potential loss to the bank if a borrower fails to meet its obligations (interest, principal amounts). Credit risk is the single largest risk banks face (Apostolik et al. 2009). The Basel Accord allows banks to take the internal ratings-based approach for credit risk. Banks can internally develop their own credit risk models for calculating expected loss. The key risk parameters to be estimated are probability of default (PD), loss given default (LGD) and exposure at default (EAD). Expected Loss = $PD \times LGD \times EAD$ (Basel Committee on Banking Supervision 2005a, 2005b).

Liquidity risk, treated separately from the other risks, takes two forms—asset liquidity risk and funding liquidity risk. A bank is exposed to asset-liquidity risk when a transaction cannot be executed at the prevailing market prices, which could be a consequence of the size of the position relative to the normal trading lot size. Funding liquidity risk refers to the inability to meet cash flow obligations, and is also known as cash flow risk (Jorion 2007). Banks are required to establish a robust liquidity risk management framework that would ensure sufficient liquidity is maintained, including the ability to withstand a range of stress events. A sound process for the identification, measurement, monitoring and control of liquidity risk should be implemented (Basel Committee on Banking Supervision 2008).

Operational risk is defined by BCBS as the risk of loss resulting from "inadequate or failed internal processes, people and systems or from external events" and is a "fundamental element of risk management" at banks. This definition includes legal risk, but excludes strategic and reputational risk. It is considered inherent in all banking products, activities, processes and systems (Basel Committee on Banking Supervision 2011). In the annual reports, operational risk was varyingly presented and included a number of sub risks, and could be referred to more as non-financial risk. It included, among many others, fraud risk, cyber security, clients products and business practices, information and resiliency risk, money laundering and financial crime risks, vendor and outsourcing risks, technology risk, business disruption risks. In some instances, banks have reported compliance and legal risk also under operational risk.

To determine the risks specific to banks, as an alternate to leveraging the existing literature, a review was done of bank annual reports. Based on the review, a taxonomy was charted of the various risk types that banks typically seek to manage as part of their business and the methodologies and tools in use. The annual reports of 10 leading banks were reviewed to determine which risk areas were specifically being reported on by these banks. The review also included identifying the specific tools, methodologies or risk management framework components that were in use. To get wider coverage, the list of banks included a representative from each region—US, predominantly globally operating banks, European banks, and also an Asian bank. Also, these banks operated a wide ranging of banking business lines—investment banking, securities trading, consumer or retail banking and corporate banking. While there were differences in the way the risks were discussed and presented, including sub risks, the top risks were largely the same and included credit risk management, market risk management, liquidity risk, operational risk.

A chart (Figure 1) depicting the taxonomy of the various risk types discussed in bank annual reports and also the various methodologies or tools (Figure 2) implemented to manage these risks is included below.
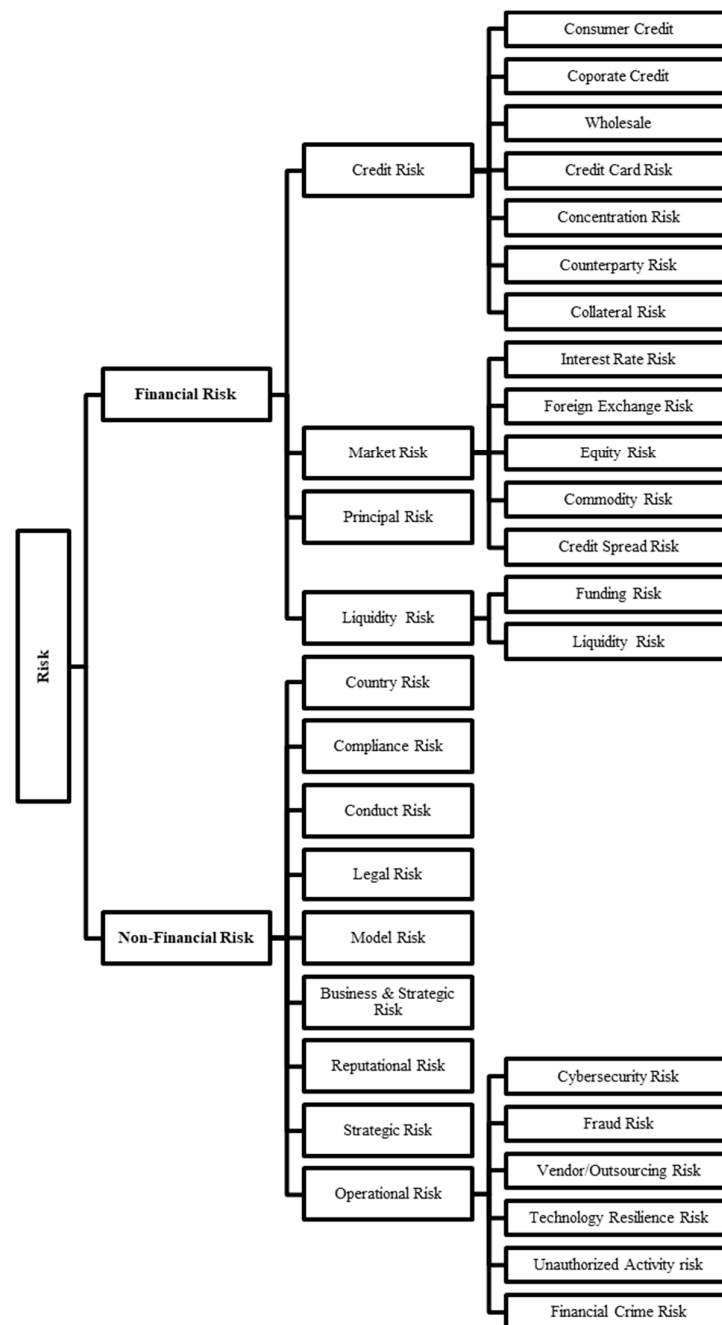
Figure 1. Taxonomy of risks.

The chief risk officer has access to risk insight and intelligence that was more retrospective in nature, such as incident analyses focusing on understanding what happened and why. Now, increasingly, they are gearing up with tools that allow for a look ahead that facilitates the predicting of potential risk incidents. Data mining, scenario modelling and forecasting are built-in features of most risk management solutions. Cognitive (pattern recognition by visualising and identifying apparent and later trends in historical data) and algorithmic (establishing causal relationships between diverse events and data sets) intelligence is making way for augmented (natural language processing and machine learning) and assistive (contextual virtual intelligent assistance) intelligence that augments and accelerates decision making (MetricStream 2018).

	Market Risk	Credit Risk	Liquidity Risk	Non-Financial Risk (Operational Risk)
Risk Management Tools				
Risk Limits	√	√	√	
Credit Risk limits		√		
Value at Risk	√			
Earnings at Risk	√			
Expected Shortfall	√			
Economic Value Stress Testing	√			
Economic Capital	√	√	√	√
Risk Sensitivities	√			
Risk Assessment (RCSA)				√
Operational Risk Losses				√
Loss Distribution Approach				√
Scenario Analysis	√	√	√	√
Tail Risk Capture	√	√	√	√
Stress Testing	√	√	√	√
Scoring Models		√		
Rating Models		√		
Exposure - Probability of Default - Loss Given Default - Exposure at Default		√		
Back Testing	√	√	√	
Risk Management Framework Components				
Risk Appetite	√	√	√	√
Risk Identification	√	√	√	√
Risk Assessment	√	√	√	√
Risk Measurement	√	√	√	√
Risk Testing	√	√	√	√
Risk Monitoring	√	√	√	√
Risk reporting	√	√	√	√
Risk Oversight	√	√	√	√
Capital Management (calculation and allocation) - CCAR - ICAAP	√	√	√	√

Figure 2. Risk Management Methods and Tools.

2.2. Machine Learning

Machine learning has been explained as lying at the intersection of computer science, engineering and statistics. It has been highlighted as a tool that can be applied to various problems, especially in fields that require data to be interpreted and acted upon (Awad and Khanna 2015). Machine learning delivers the capability to detect meaningful patterns in data, and has become a common tool for almost any task faced with the requirement of extracting meaningful information from data sets. When faced with the requirement of extracting meaningful information from data, and the consequent complexity of patterns to be studied, a programmer may not be able to provide explicit and detailed specification on the execution process. Machine learning addresses this challenge by "endowing programs" with the ability to "learn and adapt". The machine learning programs learn and improve, and can be

applied when the problem that has to be dealt has the dual challenge of complexity and the need for adaptability (Shalev-Shwartz and Ben-David 2014).

Machine learning tools that are driving the advances in search engines and self-driving cars can be adopted and applied to the financial sector. A variety of technological developments have contributed to the financial sector being able to explore and mine a voluminous data infrastructure that includes diverse sets of unstructured forms of financial data about markets and consumers. Economists are increasingly adopting machine learning, in conjunction with other tools and expertise to evaluate complex relationships, despite machine learning's limitations in being able to determine causality. The adoption of machine learning has been motivated by the potential opportunities for cost reduction, improved productivity and improved risk management. New regulations have also pushed the banks to automate with the need to have efficient regulatory compliance (Financial Stability Board 2017).

Data driven and computational-based, machine learning algorithms rely less on assumptions about the data, including about the distribution. While they are considered more robust and better at addressing complex non-linear relationships, they also are seen as being difficult to interpret (Galindo and Tamayo 2000).

Recent years have seen a surge in the amount of data gathered within financial institutions (FI). A big push towards the digitalisation of services and increased regulatory reporting requirements has resulted in a large amount of unstructured data being created and/or collected at a high frequency. This data comes from various sources, including consumer apps, client interactions, metadata and other external data sources. The desire to enhance their analytical capabilities and automate across business lines, including risk management, by managing and mining these increased volumes and a variety of data has led financial institutions to explore powerful and analytical solutions, a consequence of which is the rise in interest and the popularity of machine learning and artificial intelligence within the FI community (Van Liebergen 2017). Machine learning is widely seen in the financial services sector as having the potential to deliver the analytical capability that FIs desire. Machine learning is capable of impacting every aspect of the FI's business model—improving insight into client preferences, risk management, fraud detection, conduct monitoring, client support automation and even automated identity verification when coupled with biometrics.

Van Liebergen (2017) introduces the field, and through discussions with the Institute of International Finance and technology ventures, explains use cases within financial institutions. He discusses applications in the area of credit risk modelling, detection of credit card fraud and money laundering and surveillance of conduct breaches at FIs. He also highlights that Machine learning seeks to predict "out-of-sample" while learning "found in-sample" (past) correlations, while falling short of providing an explanation for the analysed relationship. This could create complexities around model development and evaluation.

Machine learning also plays a role at the Securities and Exchange Commission (SEC) in the risk assessment process in identifying misconduct. While this is applicable from a supervisory perspective and for the oversight of systemic risks, it can also serve as a guide for a bank on how similar machine learning techniques can be applied in risk assessments for the detection of misconduct (internal or external) including risk assessments on corporate issuers or counterparties (Bauguess 2015). In computational finance, machine learning has great potential and could be variedly used, ranging from the comprehensive exploratory data analysis to the presentation/visualisation of modelling results (Kanevski and Timonin 2010).

Some of the cons of machine learning, as argued, are that they are more "black box" in nature, with results at times being difficult to interpret. It is argued that they are also sensitive to outliers, resulting in the overfitting of the data and counterintuitive predictions. They are also argued to have the pros of being able to be a better fit for non-linear relationships between the explanatory variables and explained variables, and also that the ability for them to apply a broader set of variables tends to improve accuracy (Bacham and Zhao 2017).

3. Materials and Methods

To carry out the review of literature that researches the application of machine learning in bank risk management, two sets of key words were used in the search for related papers. The search for papers was done using the scholar.google.com, SSRN and ProQuest databases. The search was largely focused on papers after 2007 to capture developments since the global financial crisis; however papers prior to that period were also included if they were referenced in other recent papers.

The first group of words was 'machine learning', in line with the topic. The second group comprised terms that were identified from the review of the bank annual reports. This includes risk types, as listed in the risk taxonomy and risk management tools or methods that were identified from the bank annual reports. Taxonomy is as shown in Figure 1, and methods as in Figure 2.

The search and review was limited to conference papers, journal articles and selected theses (post graduate or doctoral). The review has not considered articles, white papers, vendor papers or web articles that have just made reference to machine learning without providing details on how, or that made references to the application of any specific algorithm, though many such articles did come up in the search. In particular, there are a large number of articles, web and magazines, and publications that include machine learning as a solution or as a generic and general recommendation without providing further details on how a given specific problem can be addressed.

The review has looked at only papers that have analysed the topic with a level of depth, namely, by making references to specific algorithms or providing a design or model for how ML can be implemented. Articles or papers or conference proceedings that have made only a cursory or a general reference to the application of ML in the risk management space have not been considered for this research. It is noted that there are many references available where the authors or speakers have proposed that ML or AI can be applied in the management of risk; however, many of them stop short of providing clarity on *which* algorithms, or fail to provide examples of how ML/AI has been applied in a test or industry setup.

The methodological framework for this research was determined by analysing the various problem areas related to machine learning and risk management in banks. The articles were classified to understand: (i) the risk area they focused on; (ii) the risk management tool or risk management framework component they targeted; or (iii) the algorithms that were applied/studied/proposed. The survey was also seeking to review papers that focused more on risk assessment and measurement.

Risk areas such as cybersecurity and fraud risk have been dealt with widely; however, the focus in this review has been only on cases where they specifically relate to banking risk management use cases. Papers that focus the research on operational matters, such as credit risk management solutions that address the operational process of credit review and approval, or tools that are focused on supporting traders and trading risk managers in the order and trade management process, have not been considered. Additionally, operational risk management solutions that fit within the operational process to mitigate operational events/incidents (e.g., robotics process automation, STP, anomaly detection) have not been researched.

An overview of the papers that were reviewed is included in Appendix A.

3.1. Credit Risk

The assessment of credit risk remains an important and challenging research topic in the field of finance, with initial efforts dating back to the last century. On the back of the global financial crisis events and the consequent increased regulatory focus, the credit risk assessment process has seen an increased interest within the academic and business community. The general approach to credit risk assessment has been to apply a classification technique on past customer data, including on delinquent customers, to analyse and evaluate the relation between the characteristics of a customer and their potential failure. This could be used to determine classifiers that can be applied in the categorisation of new applicants or existing customers as good or bad (Wang et al. 2005).

Credit risk evaluation occupies an important place within risk management. Techniques such as Logistic regression and discriminant analysis are traditionally used in credit scoring to determine likelihood of default. Support Vector machines are successful in classifying credit card customers who default. They were also found to be competitive in discovering features that are most significant in determining risk of default when tested and compared against the traditional techniques (Bellotti and Crook 2009). Credit risk modelling for the calculation of credit loss exposure involves the estimation of the Probability of Default (PD), the Exposure at Default (EAD) and the Loss Given Default (LGD). This is emphasised by the Basel II accord. Predominant methods to develop models for PD are classification and survival analysis, with the latter involving the estimation of whether the customer would default and when the default could occur. Classifier algorithms were found to perform significantly more accurately than standard logistic regression in credit scoring. Also, advanced methods were found to perform extremely well on credit scoring data sets such as artificial neural networks, performing better than extreme learning machine (Lessmann et al. 2015).

Through the Basel accord requirements, the need to allocate capital in an efficient and profitable manner has lead FIs to build credit scoring models to assess the default risk of their customers. Again, SVM has been shown to yield significantly better results in credit scoring (Van Gestel et al. 2003). An accurate prediction of estimated probability of default delivers more value to risk management in comparison to a binary classification of clients as either credible or not-credible. A number of techniques are used in credit scoring, such as discriminant analysis, logistic regression, Bayes classifier, nearest neighbour, artificial neural networks and classification trees. Artificial neural networks have been shown to perform classifications more accurately than the other five methods (Yeh and Lien 2009)

Methods and models are being constantly developed to address a significant issue at banks, namely, the correct classification of customers and the estimation of credit risk. The various approaches applied in these methods seek to increase the accuracy of creditworthiness predictions that could lead to a bigger and profitable loan portfolio. Neural networks have proven to be of significant value in the credit risk decision process, and their application in company distress predictions was reported to be beneficial in credit risk evaluation (Wójcicka 2017).

While credit risk is the most researched and evaluated risk area for the application of machine learning, this is not a new phenomenon. Dating as far back as 1994, Altman and colleagues conducted an analysis comparing traditional statistical methods of distress and bankruptcy prediction with alternative neural network algorithm, and concluded that a combined approach of the two improved accuracy significantly (Aziz and Dowling 2018).

Hand and Henley (1997) argued that "credit scoring is the term used to describe formal statistical methods which are used for classifying applicants for credit into "good" and "bad" risk classes". Credit scoring models are multivariate statistical models applied to economic and financial indicators to predict the default risk of individuals or companies. These indicators are assigned a weight relative of importance in predictions, and are fed as input to arrive at an index of creditworthiness. This numerical score serves as a measure of the borrower's probability of default. The support vector machine technique was concluded as being the most widely applied in credit risk evaluations. Hybrid SVM models have been proposed to improve the performance by adding methods for the reduction in the feature subset. These, however, only classify, and don't provide an estimation of the probability of default (Keramati and Yousefi 2011).

The dramatic growth in consumer credit has increased the importance of credit scoring models. The bulk of the research appears to be focussed on credit scoring techniques, as seen in the number of papers focussing on this area (Ala'raj and Abbod 2016a; Ala'Raj and Abbod 2016b; Bellotti and Crook 2009; Cao et al. 2013; Van Gestel et al. 2003; Guegan et al. 2018; Huang et al. 2007; Keramati and Yousefi 2011; Lai et al. 2006; Lessmann et al. 2015; Van-Sang and Nguyen 2016; Malhotra and Malhotra 2003; Wang et al. 2015; Wójcicka 2017). Predominantly, the focus is on classification and the application of algorithms that enable this. A number of papers evaluate the various algorithms in an attempt to identify the most efficient and accurate prediction algorithm. The papers make a

case that machine learning delivers comparable accuracy and is better equipped to capture non-linear relationships common to credit risk (Bacham and Zhao 2017; Hamori et al. 2018; Zhang 2017).

Zhou and Wang (2012) propose allocating weights to decision trees for better prediction. They put forward an improved random forest algorithm for predictions. The algorithm, during aggregation, allocates weights which are calculated based on out-of-bag errors in training to the decision trees in the forest. They attempt to address the binary classification problem, and their experiment shows that the proposed algorithm beats the original random forest and other popular classification algorithms (SVM, KMM, C4.5) in terms of balanced and overall accuracy metrics.

Some papers focus on a comparison against traditional statistical methods to highlight the efficiency in applying machine learning algorithms. Galindo and Tamayo (2000) research, through a comparative analysis of statistical and machine learning classification techniques, the credit portfolios of institutions to find accurate predictions of individual risk. They built more than 9000 models as part of the study and ranked the performance of the various algorithms. They show that the CART decision tree models provided the best estimates for default, with neural networks coming second. Hamori et al. (2018) studied and compared the prediction accuracy and classification ability of bagging, random forest, boosting with neural network methods in analysing default payment data. They found boosting to be superior among the studied machine learning methods.

A number of researchers have also evaluated the application of hybrid techniques and ensemble methods to study credit scoring (Bastos 2014; Hamori et al. 2018; Raei et al. 2016). In a hybrid system, one technique is employed for the final prediction after the use of several heterogeneous techniques in the analysis (Chen et al. 2016). In dealing with credit scoring problems, ensemble learning, using regularised logistic regression, can be applied. A method of applying clustering and bagging algorithms to balance and diversify the data, followed by lasso-logistic regression ensemble to evaluate credit risks, was found to outperform many popular credit-scoring models (Wang et al. 2015).

Khandani et al. (2010), to improve classification rates of credit card holder delinquencies and defaults, constructed a nonlinear, non-parametric forecast model. The consumer credit risk model was able to identify subtle non-linear relationships in massive datasets. These relationships were typically reportedly difficult to detect when using standard consumer credit-default models such as logit, discriminant analysis or credit scores. This allows for credit line risk management, the forecasting of aggregate consumer credit delinquencies and the forecasting of consumer credit cycle.

Yu et al. (2016) propose a novel multistage deep belief network based extreme machine learning as promising tool for credit risk assessment. The framework of multistage ensemble learning paradigms, working at three stages, is shown to outperform typical single classification techniques and similar multistage ensemble learning paradigms with high prediction accuracy.

"Support Vector Machine" (SVM) is a supervised machine-learning algorithm, and while it is widely used in classification problems, it is relatively new to credit scoring. In this algorithm, each data item is plot as a point in n-dimensional space, the value of each feature is the value of a particular coordinate (n—is number of features). Classification is performed by finding the hyper-plane that is the frontier that segregates two classes (Ray 2015). The SVM has been applied as is or in some varied form to design a credit risk evaluation and credit scoring models (Bellotti and Crook 2009; Cao et al. 2013; Van Gestel et al. 2003; Huang et al. 2007; Lai et al. 2006). Harris (2013) compares SVM-based credit scoring models using broad (<90 days past due) and narrow (>90 days past due), the latter being the more traditional approach. It was found that models built using a broader definition were more accurate, allowing for improvements in prediction accuracy.

Wang et al. (2005) propose a new "fuzzy support vector machine". The algorithm seeks to discriminate good creditors from bad ones through more generalisation while preserving the ability of the fuzzy SVM to be insensitive to outliers. They present a bilateral weighted fuzzy SVM with results that show promising application in credit analysis. Huang et al. (2007) constructed a credit scoring model to evaluate an applicant's credit score from input features based on a hybrid SVM constructed using three strategies.

Yeh and Lien (2009), in their paper, have acknowledged that forecasting the probability of default (PD) is a challenge facing practitioners and researchers, and it needs more study. A few papers have the objective of going beyond just classification, so as to predict the probability of default (PD) or recovery rates (RR) (Bastos 2014; Raei et al. 2016). ThelLeast squares vector machine technique, when incorporated into a two-state model, followed by a regression step, for predicting recovery rates, was also reported to show improvement compared to traditional statistical regression models (Yao et al. 2017). Support Vector regression techniques could also be applied to increase the predictive ability of loss given default for corporate bonds outperforming statistical models (Yao et al. 2015). These papers stop short of being able to provide a quantitative value, as they seem to approach more from a classification perspective.

Raei et al. (2016) research a new hybrid model for estimating the probability of default of corporate customers in a commercial bank. They present the hybrid mode as one that can address the 'black box'—model obtained not understandable in terms of parameters—criticism of neural networks. The research combines a two-stage approach, i.e., combining comprehensibility of logit models with the predictive power of non-linear techniques like neural networks. The overall accuracy of this hybrid model was shown outperform both the base models. Low default portfolios (LDP) are those that are considered as very low risk. LDPs have a class imbalance problem as, in a class of defaulters, they are contained in smaller numbers than in a class of good payers. Gradient boosting and random forest classifiers were found to perform well in dealing with samples that exhibited a class imbalance problem (Brown and Mues 2012).

Banks seek to develop efficient models that can assess the likelihood of counterparty defaults. Barboza et al. (2017) test machine learning models to predict bankruptcy one year prior to the event comparing the performance with results from traditional methods. They report significant predictive accuracy being achieved, and also suggest that ML techniques can easily be applied for substantial classification accuracy in comparison to traditional mechanisms. Despite the concerns around the explanatory ability of the model, given the complexity of bankruptcy models, machine learning could prove to be an important aid. Yang et al. (2011) also explore a novel method to predict bankruptcy, proposing a combined method of partial least squares (PLS) -based feature selection with SVM for information fusion. A bank could benefit from the model's ability to select the financial indicators which are most relevant in the prediction process, and also from its high level of prediction accuracy.

There are also a few papers that research the area of stress testing in credit risk management (Islam et al. 2013). Stress testing requires the modelling of the link between macro-economic developments and banking variables to determine the impact of extreme scenarios on a bank. More frequently, bottom-up approaches are used where predictions about future profits/losses are made on mostly disaggregated portfolio levels, making it data intensive and difficult to identify the exact drives of loses. Predictions on an aggregated portfolio using a top down method can complement this process. A supervised learning algorithm that does not need a pre-specified model is the Least Absolute Shrinkage and Selection Operator (Lasso) method. A more involved version of the Lasso is Adaptive Lasso, which possesses attractive convergence properties. Adaptive Lasso can be used in the absence of theoretical models, as in the case of top-down stress testing, to discover a parsimonious top-down model from a set of thousand possible specifications. It was shown to give sparse, approximately unbiased solutions, by searching for variables that describe the behaviour of credit loss rates best resulting in a parsimonious description of the relation between macro-economy and credit loss rates. A key issue is the need for substantial amounts of data to train a model (Blom 2015).

Model selection and forecasting have become a challenge as stress scenarios become more comprehensive, encompassing an increasing number of primary variables. Machine learning techniques for identifying patters and relationships between data can facilitate model selection and forecasting. These techniques don't seem to be widely applied in stress testing. When there are a large number of potential covariates and the number of observations is small, lasso regressions are found to be suited to building forecasting models. They are likely to outperform traditional statistical

models in forecasting the performance indicators required in applied stress testing. An advantage of the Lasso-type estimators is that they can handle complications arising from the high dimensional nature of stress tests (Chan-Lau 2017). The Multivariate Adaptive Regression Splines (MARS), a machine learning technique, can be viewed as a generalisation of stepwise linear regression of the classification and regression tree (CART) method. Stress testing statistical regression models such as Vector Autoregression (VAR) is a common modelling approach which is known to be unable to explain the phenomenon of fat-tailed distributions. An empirical test of these models found that the MARS model exhibited greater accuracy in model testing and superior out-of-sample performance, with MARS producing more reasonable forecasts (Jacobs 2018). Probabilistic graphs may be used for modelling and assessments of credit concentration risk with a tree-augmented Bayesian network providing a better understanding of the risk. This was also found to be suitable for stress testing analyses, with the ability to provide estimates of risk of losses consequent to changes in a borrowers financial condition (Pavlenko and Chernyak 2009). EMA workbench is a software toolbox, developed by a team at Technische Universiteit Delft, TBM Faculty (Policy Analysis Section). Particular machine learning algorithms and advanced visualisation tools are used to perform multiple experiments and analyse the results, providing the ability to explore possible uncertainties and identify causes based on the inputs.

Neural networks, Support Vector Models and Random Forest appear to be the most researched algorithms in the credit risk management area.

3.2. Market Risk

Risk can be measured by the standard deviation of unexpected outcomes, also called volatility. Value at Risk (VAR) calculates the worst loss over a target horizon that will not be exceeded with a given level of confidence and captures the combined effect of underlying volatility and exposure to financial risks (Jorion 2007). Volatility forecasting in the financial markets is important in the areas of risk management and asset pricing, among others. By using NN models, the performance of the volatility estimation method can be improved (Monfared and Enke 2014).

Zhang et al. (2017) propose a model that is based on the Generalized Autoregressive Conditional Heteroskedastic (GARCH) model and Extreme Machine Learning (ELM) algorithm to estimate volatility. The model predicts the volatility of target time series using GELM-RBF and extrapolating the predicted volatilities allows for the calculation of VaR with improved performance in terms of accuracy and efficiency. The model utilises a stochastic mapping method that doesn't require the Gaussian likelihood for estimation and is a not linear data driver model.

Market risk also includes interest rate and equity risk. Interest rate curves, which is the relation between the interest rate and the time to maturity of the debt for a given borrower in a given currency, is widely used in financial engineering and market risk management. A clustering method called the "Gaussian Mixture Model" can be used to develop nonlinear models of the evolution of the parameters and then to forecast interest rate curves. This can allow for better visualisation of interest rates (Kanevski and Timonin 2010). Machine learning clustering methods designed to address Stochastic Differential Equation (SDE) can be applied to develop anticipatable VAR models that aim at being a leading risk measure of market regime change. This can partially address some of the complexity introduced by the challenging regulatory environment, such as scenario coherence (Mahdavi-Damghani and Roberts 2017).

3.3. Liquidity Risk

A number of liquidity risk problems can be solved through the use of machine learning. Measurement of liquidity risk, the analysis of key factors including the study of the interconnections between the factors can be achieved through the use of machine learning. For the purposes of estimating a risk measure Artificial Neural Networks (ANN), a genetic algorithm can be applied. ANN can be used in the approximation of the general risk trend and determination of the most influential

factors. The probability that a liquidity risk event will occur can be estimated by the application of Bayesian Networks. The ANN and the BN implementations were capable of distinguishing the most critical liquidity risk factors measuring the risk by a functional approximation and a distributional estimation, respectively (Tavana et al. 2018).

3.4. Operational Risk

Machine learning is also applied in operational areas that enable the mitigation of risk, i.e., detection and/or prevention of risks. In the area of operational risk, aside from cyber security cases, machine learning is predominantly focused on problems related to fraud detection and suspicious transactions detection.

Khrestina et al. (2017), in their paper, propose a prototype for the generation of a report that allows for the detection of suspicious transactions. The prototype uses a logistical regression algorithm. It is noteworthy that they have also included a survey of six software solutions that are currently implemented at various banks for the automation of suspicious transaction detection and monitoring processes. While the authors make a reference to algorithms, it is unclear whether these products apply machine learning techniques, and if so, with which algorithms. No further research was done on these products as this was not in the scope of the paper.

One such area where an intelligent system based on machine learning is known to add value is in the defence against spammers where the attackers' techniques evolve. Losses from spam potentially include lost productivity, disrupted communications, malware attacks and theft of data, including financial loss. Proofpoint's MLX technology uses advanced machine learning techniques to provide comprehensive spam detection that guards against the threat of spam. Millions of messages can be analysed by the technology, which also automatically refines the detection algorithm to identify and detect newer threats (Proofpoint 2010). While it was not in scope for this research, being more of an operational control to manage risk, it has been highlighted as a case of how machine learning is used in managing cybersecurity risks.

In money laundering, criminals route money through various transactions, layering them with legitimate transactions to conceal the true source of the funds. The funds typically originate from criminal or illegal activities, and can be further used in other illegal activities including the financing of terrorist activities. There has been extensive research on detecting financial crimes using traditional statistical methods, and more recently, using machine-learning techniques. Clustering algorithms identify customers with similar behavioural patterns and can help to find groups of people working together to commit money laundering (Sudjianto et al. 2010). A major challenge for banks, given the large volume of transactions per day and the non-uniform nature of many, is to be able to sort through all the transactions and identify those that are of suspicious nature. Financial institutions utilise anti-money laundering systems to filter and classify transactions based on degrees of suspiciousness. Structured processes and intelligent systems are required to enable the detection of these money laundering transactions (Kannan and Somasundaram 2017).

Money laundering is another area that poses a significant challenge to financial institutions, given the high volumes and complexity of transactions coupled with the dynamic and fast evolving nature of financial crimes, and the need to do so on real-time data sets. In the area of financial crime detection, there has been a significant amount of research in the application of statistical learning and data mining for developing classification models to flag suspicious transactions. A C5.0 algorithm was applied to predict risk levels based on the different customer potential risk factors to create the set of rules for cluster allocation. The key factors were used to characterise transaction profiles. The model reportedly provided a 99.6% correct classification rate on the test data. The number of alerted cases was reported to have reduced from the close to 30% of transactions to less than 1% (Villalobos and Silva 2017).

Credit card fraud is significantly increasing annually, costing consumers and the industry billions of dollars. Fraudsters are constantly finding newer techniques to perpetrate the crime. In order to manage the increasing fraud risk and minimise losses, banks have fraud detection systems in place.

The systems are oriented towards increasing the detection rate while minimising the false positive rate. Models are estimated based on samples of fraudulent and legitimate transactions in supervised detection methods while in unsupervised detection methods outliers or unusual transactions are identified as potential cases of fraud. Both seek to predict the probability of fraud in a given transaction. Some reported challenges in credit card fraud detection are the non-availability of real data sets, unbalanced data sets, size of the data sets and the dynamic behaviour of fraudster. Bayesian algorithms, K-Nearest neighbor, Support Vector machines (SVM) and bagging ensemble classifier based on decision tree have been varyingly used in fraud detection systems. A comparative evaluation showed that the bagging ensemble classifier based on decision tree algorithms works well, as it is independent of attribute values, and is also able to handle class imbalance (Zareapoor and Shamsolmoali 2015). False alarms, namely transactions labelled as fraudulent that are in fact legitimate, are significant, causing concerns for customers and delaying the detection of actual fraud transactions. Large Canadian banks rely heavily on NN scores, ranging from 1 to 999, with 1 being the lowest chance of a fraudulent transaction, determined by neural network algorithms. Reportedly, 20% of transactions with a NN score greater than or equal to 990 are fraudulent, causing fraud analysts to inefficiently spend time investigating legitimate transactions. A meta-classifier (a multiple algorithm learning technique) applied to a post-neural network was shown to provide quantifiable savings improvements with a larger percentage of fraudulent transactions being caught (Pun and Lawryshyn 2012).

Also, in the areas of operational risk, there are a few papers on fraud risk detection in credit cards and online banking. They concern credit card fraud detection in domains not specifically related to bank risk management or the banking industry. One would note that the algorithms they refer to were SVM, KNN, Naïve Bayes Classifier, Bagging ensemble classifier based on decision tree (Dal Pozzolo 2015; Pun and Lawryshyn 2012; Vaidya and Mohod 2014).

4. Discussion

Credit scoring involves the assignment of a numerical value to the firm (or client) indicating whether the firm (or client) is likely to default or not. Most of the research has been focused on addressing this area by handling it as a classification problem, predicting a prospective customer as "good" or "bad" so as to facilitate the credit decision and management of credit risk. There is therefore a dominance of classification related algorithms. Many papers have gone beyond only providing a classification of creditors, and have addressed how to predict of the probability of default or recovery rate i.e., estimating the probability of default (PD), loss given default (LGD), exposure at default (EAD). It would be of significant value to banks and their risk management functions to have research and models for predicting or estimating the PD, LGD, EAD, and therefore, estimating the credit loss exposure.

Machine learning techniques have been proven to perform better than traditional statistical techniques, both in classification and also predictive accuracy. The support vector machine is seen to be a widely tested and proven machine learning approach. Much empirical work is based on observational data. Selection mechanisms could result in non-random samples, either due to sample design or the behaviour of the sampled units. In the former, data are usually missing, and in the latter, there is a self-selection of the sample units. This leads to sample selection bias (Arezzo and Guagnano 2018). When the sample data being studied has a proportional representation of certain dependent variable outcomes (e.g., default, fraudulent transaction) different from their proportional representation in the population they are drawn from, it is said to be 'choice based'. This 'choice based' sampling induces a bias in the estimation (Greene 1992). As machine learning bases much of the modelling upon learning from available data, it could be prone to the same problems and biases that affect traditional statistical methods. As machine-learning methods are compared to traditional statistical techniques, it would be beneficial to evaluate and understand how problems inherent to traditional statistical research methods fare when treated by machine learning techniques.

The application of machine learning for market risk management also doesn't appear to have been adequately studied. Many papers have researched market volatility or market risk from a portfolio or investment risk management perspective. However, from a bank risk management perspective, the papers appear limited. Further research on this, especially from the point of view of stress testing or tail end risk capture, is merited. Liquidity risk that has, since the financial crisis, attracted a lot of attention from regulators, has a few use cases researched.

In the operational risk area, studies have been predominantly focused on fraud and suspicious transaction detection—problems that are typically addressed by classification algorithms. Clustering analysis, Bayesian networks, decision, classification trees, SVM are commonly noted in the application of machine learning algorithms. Neural networks have also been referred to as a very prevalent and prominent technique in credit card fraud detection.

Much of the other areas of non-financial risk management, country risk management, compliance risk management—aside from money laundering related uses—cases haven't been explored adequately. Conduct risk, which has become a key risk and assumed a high priority for regulators and banks driven by the spate of conduct issues in Europe, US and Asia Pacific also seems to be lacking in research papers, though machine learning is offered as a solution to manage conduct risk (Oliver Wyman 2017). Furthermore, many of the standard risk management tools such as risk assessments, risk monitoring and risk reporting appear as areas that could significantly benefit from further research.

Dataset sources used by researchers for their studies are varied, with some receiving data from commercial banks or from databases provided by financial services providers (e.g., Moody's), and with some researchers conducting their studies using publicly-available data. Problems highlighted are that the data can be highly skewed (e.g., low default rates), data incompleteness and data integrity (e.g., may not be labelled correctly or as expected (Khandani et al. 2010). A wider availability of real world data sets would definitely encourage more research in evaluating the many problems faced by risk management functions.

5. Conclusions

The future of machine learning in the banking and financial industry is well recognised, and it is expected that the field of risk management will also seek to apply machine learning techniques to enhance their capabilities. Despite being critiqued for operating like a black box, the ability of machine learning techniques to analyse volumes of data without being constrained by assumptions of distribution and deliver much value in exploratory analysis, classification and predictive analytics, is significant. This offers the potential to transform the area of risk management. Machine learning, identified as one of the technologies with important implications for risk management, can enable the building of more accurate risk models by identifying complex, nonlinear patterns within large datasets. This paper has presented an assessment, analysis and evaluation of the research around the application of machine learning in risk management within the banking industry. Most of the research appears to be focused around credit risk management. One could attribute this to the fact that credit risks is considered the most significant risk to a banking organisation. More specifically from a methodology perspective, credit risk management problems researched have been around credit scoring; it would go a long way to research how machine learning can be applied to quantitative areas for better computations of credit risk exposure by predicting probabilities of default, loss given default given the many complexities and the varied factors that are involved. Market risk has seen some research with machine learning being used to forecast volatility, interest rate curves, and market regime change. Liquidity risk, despite the increased attention in the industry post the concern from Regulators, has seen limited research. Given the implications to a bank's profitability and solvency as a consequence of a liquidity risk event materialising, liquidity risk would be a very good candidate to be research extensively, more notably, research around predicting liquidity risk events in isolation or as a network of factors or events. Operational risk has also seen very limited research. Research has focused on application of machine learning to prevent or detect operational risk events; however,

there is very limited application to operational risk management practices especially in the areas of risk identification, assessments, monitoring, reporting. Given the vast amounts of operational data available (internal to a bank) machine learning could be applied in effectively developing operational risk management capabilities, which has predominantly relied on qualitative factors to measure, report and manage risk. The review has shown that areas of stress testing, tail risk capture, scenario analysis—areas that rely on predictive analysis of large volumes of data—have also seen only limited research. The advantages and disadvantages of the different machine learning techniques in solving specific risk management problems can also be further evaluated and studied to maximise the value.

The review has showed that the application of machine learning in the management of banking risks such as credit risk, market risk, operational risk and liquidity risk has been explored; however, it doesn't appear commensurate with the current industry level of focus on both risk management and machine learning. In areas of market risk, operational risk, and liquidity risk research appear lacking, and there is significant potential for further study. The application of machine learning could be further researched for some areas where analysis or modelling on volumes of data with complex and non-linear computations is required. As one of a group of topics that requires a lot of analysis of different data types to predict potential events or estimate losses, these include tail risk analysis and stress testing. Measuring and reporting technology risk is still a new area and could be further researched, especially as this risk is rising up the charts and senior managers and risk managers in bank are starting to seek more insight into what the technology risk is. As banks look to mature their enterprise risk management capabilities, it would be beneficial to study how machine learning can be applied in the aggregation of risks, and enhancing risk reporting capabilities. While areas such as conduct risk could also be researched, it is noted that these areas would benefit more from application in the operational area such, as behaviour monitoring and activity monitoring. While these go towards managing risk (risk mitigation, risk detection) at the bank, they are not the risk management systems (risk measurement, risk assessment) that are the focus of this research.

In conclusion, while there has been research on the application of machine learning in risk management over the years, it still falls short and is not on par across the various areas of risk management or risk methodologies. There still remain a large number of areas as highlighted above in bank risk management that could significantly benefit from study on how machine learning can be applied to address specific problems.

Appendix A

The table provides a list of the papers that have been reviewed classifying then by the risk type, including the risk management method/tool addressed and the algorithms referred to in the application.

Risk Type	Risk Management Method/Tool	Reference	Algorithm
Compliance Risk Management	Risk Monitoring	Mainelli and Yeandle 2006	SVM
Credit Risk Management—Concentration Risk	Stress Testing	Pavlenko and Chernyak 2009	Bayesian Networks
Credit Risk Management—Consumer Credit	Exposure (PD, LGD, EAD)	Yeh and Lien 2009	Bayesclassifier, Nearest neighbor, ANN, Classification trees
Credit Risk Management—Consumer Credit	Scoring Models	Bellotti and Crook 2009	SVM
Credit Risk Management—Consumer Credit	Scoring Models	Galindo and Tamayo 2000	CART, NN, KNN
Credit Risk Management—Consumer Credit	Scoring Models	Wang et al. 2015	Lasso logistic regression
Credit Risk Management—Consumer Credit	Scoring Models	Hamori et al. 2018	Bagging, Random Forest, Boosting
Credit Risk Management—Consumer Credit	Scoring Models	Harris 2013	SVM
Credit Risk Management—Consumer Credit	Scoring Models	Huang et al. 2007	SVM
Credit Risk Management—Consumer Credit	Scoring Models	Keramati and Yousefi 2011	NN, Bayesian Classifier, DA, Logistic Regression, KNN, Decision tree, Survival Analysis, Fuzzy Rule based system, SVM, Hybrid mode
Credit Risk Management—Consumer Credit	Scoring Models	Khandani et al. 2010	CART
Credit Risk Management—Consumer Credit	Scoring Models	Lai et al. 2006	SVM
Credit Risk Management—Consumer Credit	Scoring Models	Lessmann et al. 2015	Multiple algos assessed
Credit Risk Management—Consumer Credit	Scoring Models	Van-Sang and Nguyen 2016	Deep Learning
Credit Risk Management—Consumer Credit	Scoring Models	Yu et al. 2016	Deep belief network, Extreme Machine Learning
Credit Risk Management—Consumer Credit	Scoring Models	Wang et al. 2005	SVM, Fuzzy SVM
Credit Risk Management—Consumer Credit	Scoring Models	Zhou and Wang 2012	Random Forest
Credit Risk Management—Coporate Credit	Exposure (PD, LGD, EAD)	Bastos 2014	Bagging
Credit Risk Management—Coporate Credit	Exposure (PD, LGD, EAD)	Barboza et al. 2017	Neural Network, SVM, Boosting, Bagging, Random Forest

Risk Type	Risk Management Method/Tool	Reference	Algorithm
Credit Risk Management—Coporate Credit	Exposure (PD, LGD, EAD)	Raei et al. 2016	Neural Networks
Credit Risk Management—Coporate Credit	Exposure (PD, LGD, EAD)	Yang et al. 2011	SVM
Credit Risk Management—Coporate Credit	Exposure (PD, LGD, EAD)	Yao et al. 2017	SVR
Credit Risk Management—Coporate Credit	Scoring Models	Ala'Raj and Abbod 2016b	Multiclassifer system (MCS)—Ensemble—neural networks (NN), support vector machines (SVM), random forests (RF), decision trees (DT) and naïve Bayes (NB).
Credit Risk Management—Coporate Credit	Scoring Models	Ala'raj and Abbod 2016a	GNG, MARS
Credit Risk Management—Coporate Credit	Scoring Models	Bacham and Zhao 2017	ANN, Random Forest
Credit Risk Management—Coporate Credit	Scoring Models	Cao et al. 2013	SVM
Credit Risk Management—Coporate Credit	Scoring Models	Van Gestel et al. 2003	SVM
Credit Risk Management—Coporate Credit	Scoring Models	Guegan et al. 2018	Elastic Net, random forest, Boosting, NN
Credit Risk Management—Coporate Credit	Scoring Models	Malhotra and Malhotra 2003	NN
Credit Risk Management—Coporate Credit	Scoring Models	Wójcicka 2017	Neural networks
Credit Risk Management—Coporate Credit	Scoring Models	Zhang 2017	KNN, Random Forest
Credit Risk Management—Corporate Credit	Stress Testing	Blom 2015	Lasso regression
Credit Risk Management—Corporate Credit	Stress Testing	Chan-Lau 2017	Lasso regression
Credit Risk Management—Credit Card Risk	Exposure (PD, LGD, EAD)	Yao et al. 2017	SVM
Credit Risk Management—Cross-risk	Stress Testing	Jacobs 2018	MARS
Credit Risk Management—Wholesale	Stress Testing	Islam et al. 2013	Cluster analysis
Liquidity Risk Management—Liquidity Risk	Risk Limits	Gotoh et al. 2014	vSVM
Liquidity Risk Management—Liquidity Risk	Risk Monitoring	Sala 2011	ANN

Risk Type	Risk Management Method/Tool	Reference	Algorithm
Liquidity Risk Management—Liquidity Risk	Scoring Models	Tavana et al. 2018	ANN, Bayesian Networks
Management—Consumer Credit	Scoring Models	Brown and Mues 2012	Gradient, Boosting, Random Forest, Least Squares—SVM
Market Risk Management—Equity Risk	Value at Risk	Zhang et al. 2017	GELM
Market Risk Management—Equity Risk	Value at Risk	Mahdavi-Damghani and Roberts 2017	Cluster analysis
Market Risk Management—Equity Risk	Value at Risk	Monfared and Enke 2014	NN
Market Risk Management—Interest Rate Risk	Value at Risk	Kanevski and Timonin 2010	SOM, Gaussian Mixtures, Cluster Analysis
Operational Risk Management—Cybersecurity	Risk Assessment (RCSA)	Peters et al. 2017	Non-linear clustering method
Operational Risk Management—Fraud Risk	Operational Risk Losses	Pun and Lawryshyn 2012	Neural Networks, k-Nearest Neighbor, Naïve Bayesian, Decision Tree
Operational Risk Management—Fraud Risk	Operational Risk Losses	Sharma and Choudhury 2016	SOM
Operational Risk Management—Fraud Risk	Risk Monitoring	Ngai et al. 2011	neural networks, Bayesian belief network, decision trees
Operational Risk Management—Fraud Risk	Risk Monitoring	Sudjianto et al. 2010	SVM, Classification Trees, Ensemble Learning, CART, C4.5, Bayesian belief networks, HMM
Operational Risk Management—Money Laundering/Financial Crime	Risk Monitoring	Khrestina et al. 2017	logistic regression

References

Ala'raj, Maher, and Maysam F. Abbod. 2016a. A New Hybrid Ensemble Credit Scoring Model Based on Classifiers Consensus System Approach. *Expert Systems with Applications* 64: 36–55. [CrossRef]

Ala'Raj, Maher, and Maysam F. Abbod. 2016b. Classifiers Consensus System Approach for Credit Scoring. *Knowledge-Based Systems* 104: 89–105. [CrossRef]

Apostolik, Richard, Christopher Donohue, Peter Went, and Global Association of Risk Professionals. 2009. *Foundations of Banking Risk: An Overview of Banking, Banking Risks, and Risk-Based Banking Regulation.* New York: John Wiley.

Arezzo, Maria, and Giuseppina Guagnano. 2018. Response-Based Sampling for Binary Choice Models with Sample Selection. *Econometrics* 6: 12. [CrossRef]

Awad, Mariette, and Rahul Khanna. 2015. Machine Learning in Action: Examples. *Efficient Learning Machines.* [CrossRef]

Aziz, Saqib, and Michael M. Dowling. 2018. AI and Machine Learning for Risk Management. *SSRN Electronic Journal.* [CrossRef]

Bacham, Dinesh, and Janet Zhao. 2017. Machine Learning: Challenges and Opportunities in Credit Risk Modeling. Available online: https://www.moodysanalytics.com/risk-perspectives-magazine/managing-disruption/spotlight/machine-learning-challenges-lessons-and-opportunities-in-credit-risk-modeling (accessed on 2 April 2018).

Barboza, Flavio, Herbert Kimura, and Edward Altman. 2017. Machine learning models and bankruptcy prediction. *Expert Systems with Applications* 83: 405–17. [CrossRef]

Basel Committee on Banking Supervision. 2005a. *Guidance on Paragraph 468 of the Framework Document*. Basel: Bank for International Settlements.

Basel Committee on Banking Supervision. 2005b. *An Explanatory Note on the Basel II IRB Risk Weight Functions*. Basel: Bank for International Settlements.

Basel Committee on Banking Supervision. 2006. *Minimum Capital Requirements for Market Risk*. Basel: Bank for International Settlements.

Basel Committee on Banking Supervision. 2008. *Principles for Sound Liquidity Risk Management and Supervision*. Basel: Bank for International Settlements.

Basel Committee on Banking Supervision. 2011. *Principles for the Sound Management of Operational Risk*. Basel: Bank for International Settlements, pp. 1–27.

Bastos, João A. 2014. Ensemble Predictions of Recovery Rates. *Journal of Financial Services Research* 46: 177–93. [CrossRef]

Bauguess, Scott W. 2015. *The Hope and Limitations of Machine Learning in Market Risk Assessment*. Washington, DC: U.S. Securities and Exchange Commission.

Bellotti, Tony, and Jonathan Crook. 2009. Support Vector Machines for Credit Scoring and Discovery of Significant Features. *Expert Systems with Applications*. [CrossRef]

Blom, Tineke. 2015. Top down Stress Testing: An Application of Adaptive Lasso to Forecasting Credit Loss Rates. Master's Thesis, Faculty of Science, Hongkong, China.

Brown, Iain, and Christophe Mues. 2012. An experimental comparison of classification algorithms for imbalanced credit scoring data sets. *Expert Systems with Applications* 39: 3446–53. [CrossRef]

Cao, Jie, Hongke Lu, Weiwei Wang, and Jian Wang. 2013. A Loan Default Discrimination Model Using Cost-Sensitive Support Vector Machine Improved by PSO. *Information Technology and Management* 14: 193–204. [CrossRef]

Chan-Lau, Jorge. 2017. Lasso Regressions and Forecasting Models in Applied Stress Testing. *IMF Working Papers* 17: 1. [CrossRef]

Chen, Ning, Bernardete Ribeiro, and An Chen. 2016. Financial Credit Risk Assessment: A Recent Review. *Artificial Intelligence Review* 45: 1–23. [CrossRef]

Dal Pozzolo, Andrea. 2015. Adaptive Machine Learning for Credit Card Fraud Detection. Unpublished doctoral dissertation, Université libre de Bruxelles, Faculté des Sciences—Informatique, Bruxelles.

Deloitte University Press. 2017. *Global Risk Management Survey*, 10th ed. Deloitte University Press: Available online: https://www2.deloitte.com/tr/en/pages/risk/articles/global-risk-management-survey-10th-ed.html (accessed on 4 October 2018).

Financial Stability Board. 2017. Artificial Intelligence and Machine Learning in Financial Services. Market Developments and Financial Stability Implications. *Financial Stability Board*. November 1. Available online: http://www.fsb.org/2017/11/artificial-intelligence-and-machine-learning-in-financial-service/ (accessed on 2 July 2018).

Galindo, Jorge, and Pablo Tamayo. 2000. Credit Risk Assessment Using Statistical and Machine Learning: Basic Methodology and Risk Modeling Applications. *Computational Economics* 15: 107–43. [CrossRef]

Gotoh, Jun-ya, Akiko Takeda, and Rei Yamamoto. 2014. Interaction between financial risk measures and machine learning methods. *Computational Management Science* 11: 365–402. [CrossRef]

Greene, William H. 1992. A Statistical Model for Credit Scoring. NYU Working Paper No. EC-92-29. Available online: https://ssrn.com/abstract=1867088 (accessed on 8 April 1992).

Guegan, Dominique, Peter Addo, and Bertrand Hassani. 2018. Credit risk analysis using machine and deep learning models. *Risks* 6: 38.

Hamori, Shigeyuki, Minami Kawai, Takahiro Kume, Yuji Murakami, and Chikara Watanabe. 2018. Ensemble Learning or Deep Learning? Application to Default Risk Analysis. *Journal of Risk and Financial Management* 11: 12. [CrossRef]

Hand, David J., and William E. Henley. 1997. Statistical Classification Methods in Consumer Credit Scoring: A Review. *Journal of the Royal Statistical Society Series A: Statistics in Society*. [CrossRef]

Harris, Terry. 2013. Quantitative credit risk assessment using support vector machines: Broad versus Narrow default definitions. *Expert Systems with Applications* 40: 4404–13. [CrossRef]

Helbekkmo, Hans, Alok Kshirsagar, Andreas Schlosser, Francesco Selandari, Uwe Stegemann, and Joyce Vorholt. 2013. *Enterprise Risk Management—Shaping the Risk Revolution*. New York: McKinsey & Co., Available online: www.rmahq.org (accessed on 18 June 2018).

Huang, Cheng Lung, Mu Chen Chen, and Chieh Jen Wang. 2007. Credit Scoring with a Data Mining Approach Based on Support Vector Machines. *Expert Systems with Applications* 33: 847–56. [CrossRef]

Hull, John. 2012. *Risk Management and Financial Institutions*. New York: John Wiley and Sons, vol. 733.

Islam, Tushith, Christos Vasilopoulos, and Erik Pruyt. 2013. Stress—Testing Banks under Deep Uncertainty. Paper presented at the 31st International Conference of the System Dynamics Society, Cambridge, MA, USA, July 21–25; Available online: http://repository.tudelft.nl/islandora/object/uuid:c162de43-4235-4d29-8eed-3246df87e119?collection=education (accessed on 17 July 2018).

Jacobs, Michael, Jr. 2018. The validation of machine-learning models for the stress testing of credit risk. *Journal of Risk Management in Financial Institutions* 11: 218–43.

Jorion, Philippe. 2007. *Value at Risk: The New Benchmark for Managing Financial Risk*. New York: McGraw-Hill.

Kanevski, Mikhail F., and Vadim Timonin. 2010. Machine learning analysis and modeling of interest rate curves. Paper presented at the 18th European Symposium on Artificial Neural Networks ESANN, Bruges, Belgium, April 28–30; Available online: https://www.elen.ucl.ac.be/Proceedings/esann/esannpdf/es2010-17.pdf (accessed on 18 June 2018).

Kannan, Somasundaram, and K. Somasundaram. 2017. Autoregressive-Based Outlier Algorithm to Detect Money Laundering Activities. *Journal of Money Laundering Control* 20: 190–202. [CrossRef]

Keramati, Abbas, and Niloofar Yousefi. 2011. A proposed classification of data mining techniques in credit scoring. Paper presented at the 2011 International Conference of Industrial Engineering and Operations Management, Kuala Lumpur, Malaysia, January 22–24.

Khandani, Amir E., Adlar J. Kim, and Andrew W. Lo. 2010. Consumer credit-risk models via machine-learning algorithms. *Journal of Banking & Finance* 34: 2767–87.

Khrestina, Marina Pavlovna, Dmitry Ivanovich Dorofeev, Polina Andreevna Kachurina, Timur Rinatovich Usubaliev, and Aleksey Sergeevich Dobrotvorskiy. 2017. Development of Algorithms for Searching, Analyzing and Detecting Fraudulent Activities in the Financial Sphere. *European Research Studies Journal* 20: 484–98.

Lai, Kin Keung, Lean Yu, Ligang Zhou, and Shouyang Wang. 2006. Credit risk evaluation with least square support vector machine. In *International Conference on Rough Sets and Knowledge Technology*. Berlin/Heidelberg: Springer, pp. 490–95.

Lessmann, Stefan, Bart Baesens, Hsin Vonn Seow, and Lyn C. Thomas. 2015. Benchmarking State-of-the-Art Classification Algorithms for Credit Scoring: An Update of Research. *European Journal of Operational Research* 247: 124–36. [CrossRef]

Mahdavi-Damghani, Babak, and Stephen Roberts. 2017. A Proposed Risk Modeling Shift from the Approach of Stochastic Differential Equation towards Machine Learning Clustering: Illustration with the Concepts of Anticipative and Responsible VaR. *SSRN Electronic Journal*. [CrossRef]

Mainelli, Michael, and Mark Yeandle. 2006. Best execution compliance: New techniques for managing compliance risk. *The Journal of Risk Finance* 7: 301–12. [CrossRef]

Malhotra, Rashmi, and D. K. Malhotra. 2003. Evaluating Consumer Loans Using Neural Networks. *Omega* 31: 83–96. [CrossRef]

MetricStream. 2018. The Chief Risk Officer's Role in 2018 and Beyond Managing the Challenges and Opportunities of a Digital Era New Roles of the CRO. Available online: https://www.metricstream.com/insights/chief-risk-officer-role-2018.htm (accessed on 23 June 2018).

Monfared, Soheil Almasi, and David Enke. 2014. Volatility Forecasting Using a Hybrid GJR-GARCH Neural Network Model. *Procedia Computer Science* 36: 246–53. [CrossRef]

Ngai, Eric W. T., Yong Hu, Yiu Hing Wong, Yijun Chen, and Xin Sun. 2011. The Application of Data Mining Techniques in Financial Fraud Detection: A Classification Framework and an Academic Review of Literature. *Decision Support Systems* 50: 569. [CrossRef]

Oliver Wyman. 2017. Next Generation Risk Management. Available online: https://www.oliverwyman.com/content/dam/oliver-wyman/v2/publications/2017/aug/Next_Generation_Risk_Management_Targeting_A-Technology_Dividend.pdf (accessed on 1 May 2018).

Pavlenko, Tatjana, and Oleksandr Chernyak. 2009. Bayesian Networks for Modeling and Assessment of Credit Concentration Risks. *International Statistical Conference Prague.* Available online: http://www.czso.cz/conference2009/proceedings/data/methods/pavlenko_paper.pdf (accessed on 21 July 2018).

Peters, Gareth, Pavel V. Shevchenko, Ruben Cohen, and Diane Maurice. 2017. Statistical Machine Learning Analysis of Cyber Risk Data: Event Case Studies. Available online: https://ssrn.com/abstract=3073704 (accessed on 18 June 2018).

Proofpoint. 2010. MLX Whitepaper "Machine Learning to Beat Spam Today and Tomorrow". Available online: https://www.excelmicro.com/datasheets/Proofpoint-White-Paper-MLX-Technology.pdf (accessed on 2 May 2018).

Pun, Joseph, and Yuri Lawryshyn. 2012. Improving credit card fraud detection using a meta-classification strategy. *International Journal of Computer Applications* 56: 41–46. [CrossRef]

Raei, Reza, Mahdi Saeidi Kousha, Saeid Fallahpour, and Mohammad Fadaeinejad. 2016. A Hybrid Model for Estimating the Probability of Default of Corporate Customers. *Iranian Journal of Management Studies* 9: 651–73.

Ray, Sunil. 2015. Understanding Support Vector Machine Algorithm from Examples (along with Code). Available online: https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/ (accessed on 16 August 2018).

Sala, Jordi Petchamé. 2011. Liquidity Risk Modeling Using Artificial Neural Network. Master's thesis, Universitat Politècnica de Catalunya, Barcelona, Spain.

Saunders, Anthony, Marcia Millon Cornett, and Patricia Anne McGraw. 2006. *Financial Institutions Management: A Risk Management Approach.* New York: McGraw-Hill.

Shalev-Shwartz, Shai, and Shai Ben-David. 2014. *Understanding Machine Learning: From Theory to Algorithms.* Cambridge: Cambridge University Press. [CrossRef]

Sharma, Shashank, and Arjun Roy Choudhury. 2016. Fraud Analytics: A Survey on Bank Fraud Prediction Using Unsupervised Learning Based Approach. *International Journal of Innovation in Engineering Research and Technology* 3: 1–9.

Sudjianto, Agus, Sheela Nair, Ming Yuan, Aijun Zhang, Daniel Kern, and Fernando Cela-Díaz. 2010. Statistical Methods for Fighting Financial Crimes. *Technometrics* 52: 5–19. [CrossRef]

Tavana, Madjid, Amir Reza Abtahi, Debora Di Caprio, and Maryam Poortarigh. 2018. An Artificial Neural Network and Bayesian Network Model for Liquidity Risk Assessment in Banking. *Neurocomputing* 275: 2525–54. [CrossRef]

Vaidya, Avanti H., and Sudhir W. Mohod. 2014. Internet Banking Fraud Detection using HMM and BLAST-SSAHA Hybridization. *International Journal of Science and Research (IJSR)* 3: 574–9.

Van Gestel, Ir Tony, Bart Baesens, Ir Joao Garcia, and Peter Van Dijcke. 2003. A support vector machine approach to credit scoring. In *Forum Financier—Revue Bancaire Et Financiaire Bank En Financiewezen.* Bruxelles: Larcier, pp. 73–82. Available online: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.93.6492&rep=rep1&type=pdf (accessed on 7 July 2018).

Van Liebergen, Bart. 2017. Machine Learning: A Revolution in Risk Management and Compliance? *Journal of Financial Transformation* 45: 60–67.

Van-Sang, Ha, and Ha-Nam Nguyen. 2016. Credit Scoring with a Feature Selection Approach Based Deep Learning. In *MATEC Web of Conferences.* Les Ulis: EDP Sciences, volume 54, p. 05004. Available online: https://www.matec-conferences.org/articles/matecconf/abs/2016/17/matecconf_mimt2016_05004/matecconf_mimt2016_05004.html (accessed on 19 July 2018).

Villalobos, Miguel Agustín, and Eliud Silva. 2017. A Statistical and Machine Learning Model to Detect Money Laundering: An Application. Available online: http://hddavii.eventos.cimat.mx/sites/hddavii/files/Miguel_Villalobos.pdf (accessed on 21 June 2018).

Wang, Yongqiao, Shouyang Wang, and Kin Keung Lai. 2005. A New Fuzzy Support Vector Machine to Evaluate Credit Risk. *IEEE Transactions on Fuzzy Systems* 13: 820–31. [CrossRef]

Wang, Hong, Qingsong Xu, and Lifeng Zhou. 2015. Large Unbalanced Credit Scoring Using Lasso-Logistic Regression Ensemble. *PLoS ONE* 10: e0117844. [CrossRef] [PubMed]

Wójcicka, Aleksandra. 2017. Neural Networks vs. Discriminant Analysis in the Assessment of Default. *Electronic Economy*, 339–49. [CrossRef]

Yang, Zijiang, Wenjie You, and Guoli Ji. 2011. Using partial least squares and support vector machines for bankruptcy prediction. *Expert Systems with Applications* 38: 8336–42. [CrossRef]

Yao, Xiao, Jonathan Crook, and Galina Andreeva. 2015. Support vector regression for loss given default modelling. *European Journal of Operational Research* 240: 528–38. [CrossRef]

Yao, Xiao, Jonathan Crook, and Galina Andreeva. 2017. Enhancing two-stage modelling methodology for loss given default with support vector machines. *European Journal of Operational Research* 263: 679–89. [CrossRef]

Yeh, I. Cheng, and Chehui Lien. 2009. The Comparisons of Data Mining Techniques for the Predictive Accuracy of Probability of Default of Credit Card Clients. *Expert Systems with Applications*. [CrossRef]

Yu, Lean, Zebin Yang, and Ling Tang. 2016. A Novel Multistage Deep Belief Network Based Extreme Learning Machine Ensemble Learning Paradigm for Credit Risk Assessment. *Flexible Services and Manufacturing Journal* 28: 576–92. [CrossRef]

Zareapoor, Masoumeh, and Pourya Shamsolmoali. 2015. Application of Credit Card Fraud Detection: Based on Bagging Ensemble Classifier. *Procedia Computer Science* 48: 679–86. [CrossRef]

Zhang, Wenhao. 2017. Machine Learning Approaches to Predicting Company Bankruptcy. *Journal of Financial Risk Management* 6: 364–74. [CrossRef]

Zhang, Heng Guo, Chi Wei Su, Yan Song, Shuqi Qiu, Ran Xiao, and Fei Su. 2017. Calculating Value-at-Risk for High-Dimensional Time Series Using a Nonlinear Random Mapping Model. *Economic Modelling* 67: 355–67. [CrossRef]

Zhou, Lifeng, and Hong Wang. 2012. Loan Default Prediction on Large Imbalanced Data Using Random Forests. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. [CrossRef]