International Journal of Management, IT & Engineering
Vol. 9 Issue 8, August 2019,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

Overview of Machine Learning Applications in Financial Sector

Pallavi Mitragotri*

Vinod Kumar Pal**

Abstract

Machine Learning(ML) and Artificial Intelligence(AI) are virtually changing every aspect of our lives by having its applications in various sectors and areas like Oil and Gas, Health care, Government, Financial services, Transportation, Marketing &Sales, and customer service etc. This disruptive technology has ability to learn from data, identify patterns and make predictions with minimum human intervention. Machine learning can help financial institutions to detect and prevent frauds, predict upcoming opportunities, analyze business cases and take appropriate actions, do sentiment analysis, portfolio management etc. The purpose of this paper is to take overview of some of the applications of ML in Financial Services.

Keywords:

Machine Learning;
Financial Sector;
Artificial Intelligence;
Algorithm;

Author correspondence:

Pallavi Mitragotri
Email: Pallavi.Mitragotri@gmail.com

1. Introduction

Machine Learning is the science where computers learn from data, identify patterns and make predictions with minimum human intervention. Machine learning in its basic forms uses algorithms to parse data, learn from it and then make meaningful predictions which can be useful in decision making. Growing volumes of data and its wide varieties, complex processing and affordability are the main factors which made artificial intelligence increasingly popular. AI has made it possible to make precise computational models which can analyze complex and bigger volumes of data in real time, with more accurate results in lesser time and on a very large scale. Machine learning consists of two important parts, Data and Algorithm where data is the information you feed and algorithms are the methods you use to find relationships among data using statistical models and make predictions. The machine learning learns through these training algorithms and data scientist can retrain the models frequently to keep them updated and effective. Machine learning can be applied to statistical problems which could be Regression, clustering or classification class. Regression and classification class problems can be solved by supervised learning while the clustering problems are solved by unsupervised learning. Out of various machine learning methods, Supervised, Unsupervised, Semi-supervised and reinforcement learning are most popular and widely adopted. The purpose of the supervised learning is to establish the relationship between two datasets and to use one dataset to forecast the other. Supervised learning is commonly used in applications where historical data predicts the likely future events. E.g. detecting likely fraudulent credit card transactions based on historical transactions. Unsupervised learning is used where there are no historical labels and there is need to understand the structure of data and

*Department of Finance, Balaji Institute of International Business, Pune, (India)

**Balaji Institute of International Business, Pune, (India)

to identify main drivers behind it. E.g. understanding consumer behavior and use it in segmentation and marketing. Semi-supervised learning is used and applied to the same kind of scenarios where the Supervised learning but this technique uses both labeled and unlabeled data for training. Typically small amount of labeled data, along with large amount of unlabeled data is used as unlabeled data requires less time, money and efforts to acquire. This technique can be used with methods such as classification, regression and prediction. E.g. Face Recognition in Webcam. Reinforcement learning is used to explore and choose the successive strategies to achieve maximum output/profit. The following figure shows the various machine learning approaches.

		Linear methods	Non-linear methods
Problem type	**Supervised**		
	Regression	• Principal components • Ridge • Partial least squares • LASSO	Penalized regression: • LASSO • LARS • elastic nets Neural networks and deep learning
	Classification	Support vector machines	Decision trees: • classification trees • regression trees • random forests Support vector machines Deep learning
	Unsupervised		
	Clustering*	Clustering methods: K- and X-means, hierarchical Principal components analysis Deep learning * Since unsupervised methods do not describe a relation between a dependent and interdependent variable, they cannot be labelled linear or non-linear.	

Table 1 – Overview of machine learning methods

Table:- Overview of Machine Learning Methods(Advisor, n.d.)

Most industries working on voluminous and complex data have realized the importance of machine learning and artificial intelligence in gaining competitive advantage. Oil and Gas, Health care, Government, Financial services, Transportation, and Marketing &Sales are the some of the industries which are implementing machine learning. Large volumes of historical data and quantitative nature of finance domain makes use of machine learning inevitable and those who will not use, will be left behind. The future of financial services without machine learning is hard to imagine due to increasing no of people defaulting on loans, cybercrimes, cheating, fraudulent transactions etc. With the invention of various advance technologies and algorithms, Machine learning plays an integral part in many phases of financial ecosystem right from analyzing the credit history, approving loans to detection of frauds. The figure 1shows the growth of AI/ML market in different geographical regions over 10years which shows accelerating adoption and its critical importance.

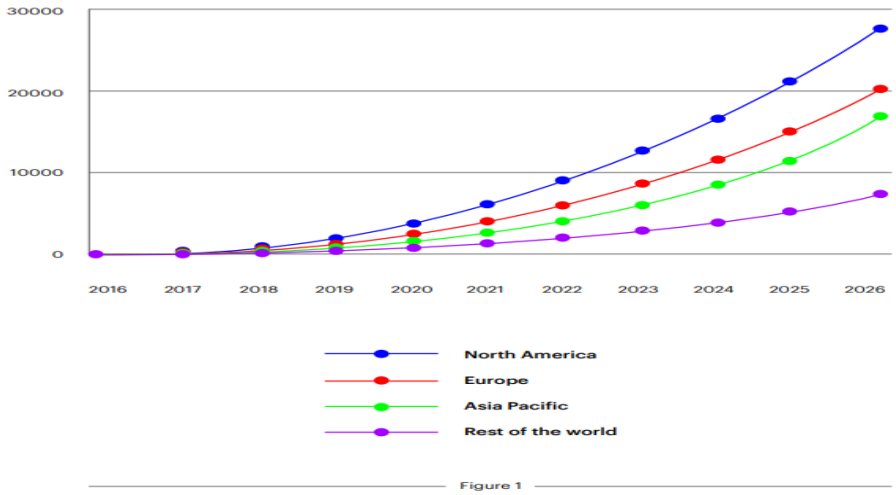

Fig 1: Growth of AL/ML market in different Regions over 10 years
Source: https://www.accenture.com/_acnmedia/PDF-84/Accenture-Machine-Leaning-Insurance.pdf

2. Applications of Machine Learning and AI in Financial Services

ML in financial services offers insights from data where in algorithms are trained to analyze unstructured data, find non-linear relationship among various variables and factors to predict the outcome, learning patterns from complex interactions to suggest appropriate actions. Following are the most widely used applications of ML in financial services.

1) Credit scoring:
Even after careful verification of credibility of loan seeking organizations, many of them are defaulting on Principle and Interest payments to the financing institutions. Lenders use credit scores to speed up the process with potentially limiting the incremental risk of default. Most credit scoring models use 'data on transactions made' and 'payments history' as the basis. These models use statistical tools such as decision trees, regression analysis and classical statistical theories. But these models are resilient only for limited amounts of structured data and when it comes to large amounts of data or unstructured data, assumptions of the classical statistical theories fail and affect the accuracy of the analysis and prediction. To improve the credit rating accuracy the lenders are increasingly using the large amounts of unstructured or semi-structured data from various other resources such as behavior on social networking platforms, online shopping behavior, their nationality, occupation, salary, experience etc. Integrated centralized system which can accurately assess the credibility of the prospective customers is critical to the lending. Application of machine learning algorithms and models to this large, unstructured/semi-structured data has enabled to summarize massive information and uncover trends that can be used for credit decision. Thus ML can help to fasten the process of assessing the borrower quality by reducing due diligence required and to make quicker credit decisions. Many traditional credit scoring models require the credit history for the specific period of time to have valid score. The potential borrower may not have the required credit history and hence may not be able to borrow the funds. Analysis of unstructured data collected from sources like social networking sites, mobile etc. can help the lenders to develop the assessment of ability and willingness to pay and arrive at the credit decision. The regression analysis uses dependent and independent variables, and use statistical processes to find out the relationship among the variables. Kount and APEX Analytics are two of the companies working in fraud detection application of machine learning. Kaunt has patented fraud prevention technology which combines supervised and unsupervised learning, device finger printing, business intelligence tools, a robust policy and rules engine and a web based case management and investigation system.

2) Fraud Detection:
With the growing number of transactions, users, and various means of doing transactions (Online, credit/debit cards, internet banking etc.), the security threats are on the rise. Traditional fraud detection techniques use the set of rules and regulations to prevent the frauds. Modern fraud detection techniques using machine learning can identify complex patterns and identify suspicious transactions which need close monitoring and investigation. Machine learning is being used to identify the non-linear relationship among various attributes and entities and to flag the complicated activities or behavioral patterns which might lead to money laundering not directly observable through individual transactions filings. E.g. Data scientist's train

machine learning models to detect a large no of micropayments and flag these transactions for potential money laundering case. But the challenge for these systems is to avoid 'the situations that are flagged' are 'not risks in reality'. Market regulators can also use ML for disclosure and risk assessment. The US SEC (Security and Exchange Commission) staff trains algorithms to understand what patterns, language or trend in underlying data may indicate possible fraud or misconduct.(Advisor, n.d.)Supervised and unsupervised learning in combination make analysis of new data and detect anomalies and possible frauds. Banks can track and monitor thousands of transactions on various parameters for each account in real time. The algorithms are trained to examine if the attempted activity is characteristic of that particular user. It the suspicious behavior is identified by the algorithm, additional information is requested from the user for authentication and validation of the transaction. If the probability of the transaction being fraud is very high then the transaction is blocked. Machine learning algorithms spot fraudulent transaction with high precision and in very less time and hence the frauds can be prevented in real time.India's cyber security market is forecasted to grow at a CAGR of over 19% during 2018-2023. As per Nasscom, Data Security Council of India & PwC Report the cyber security market has been projected to be $35 billion by 2025.The implications of security breach cost very high both in terms of finance and infrastructure. The 94-crore Cosmos Bank fraud was caused by malware attack on the bank's security system and the hackers managed to clone thousands of debit cards. Through the proxy system, 14800 fraudulent transactions were approved between 11-aug-2018 to 13-aug-2018 to withdraw 80.5Cr. and remaining 13.5 Cr were transferred online through SWIFT.

3) Portfolio Management:

In online portfolio management system using machine learning, the algorithms and statistical tools are applied to allocate, employ and optimize the clients' assets and to calibrate the financial portfolio to the objectives of the investor taking into consideration the factors such as age, income, risk taking ability and the period in which the goals should be achieved etc. the machine learning tools work on the same principles as existing analytical techniques used in traditional investing. But the ML has the advantages of ability to handle large volumes of data in real time, precision, track current trends, absence of behavioral bias and emotions which cause investors to act irrationally. The algorithms are trained to make use of historical data about the past performance of the organizations; track the news about the companies, analyze the market sentiments about the companies and draw the meaningful signs indicating the future of the company. The signals from this data are examined to track the price movements and its volatility and the algorithm calibrates to the changes in the market or to the changes of investors' goals in any. Thus the overall portfolio management process consists of three steps:1) analyze existing portfolio 2) Train algorithms to track and monitor changes 3)find the optimal portfolio selection and update portfolio.

'Blackrock' one of the world's largest investment firms offers 'Aladdin Risk Platform' which uses risk analytics and software management tools and enables the individuals and investment managers to assess the level of risk or returns in a particular portfolio. 'AXYON AI' is an Italian Fintech startup which offers the predictive modeling tool 'StocksAnalyst' which is specially designed for portfolio management. The company claims that this platform can integrate the structured and unstructured data such as financial reports, market data, news and economic indicators such as currency rates to analyze trends, draw patterns and make predictions about the said portfolio.

4) Underwriting:

Underwriting is the process of identifying the potential risk of default and covering it for a fee. Insurance, Investment banking and Commercial banking widely use underwriting. ML algorithms are trained to consider thousands of factors and learn from past claims, to prevent the possible loss and make more accurate underwriting decisions. Most insurance companies can process only 10-15% of data they have when using the traditional underwriting process and most of this data is structured. ML can be effectively used for very large amount of structured, unstructured and semi-structured data and can bring order and purpose to this data. In addition to this, ML can also be helpful in reducing the claims processing time and operational costs. Following diagram shows the potential for Machine Learning in insurance value chain.

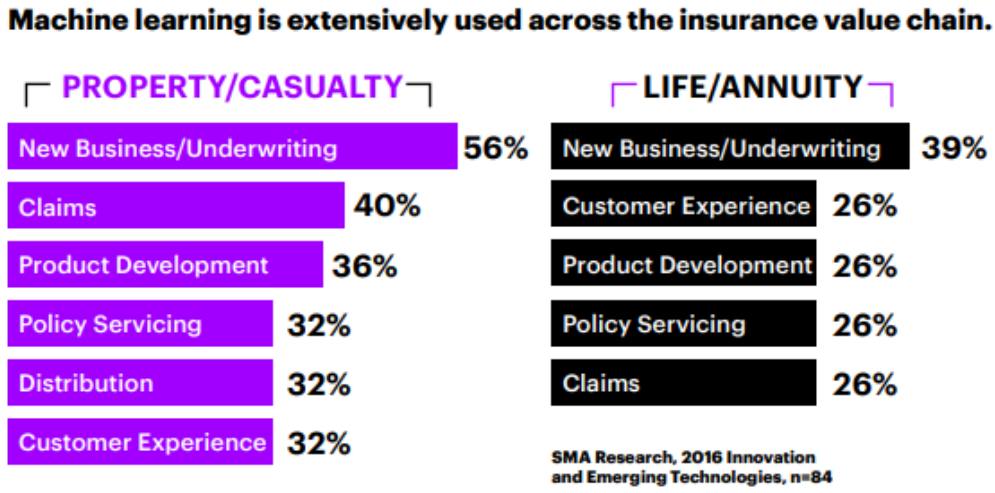

Figure 2: Insurance business areas where machine learning can be leveraged

Source: https://www.accenture.com/_acnmedia/PDF-84/Accenture-Machine-Leaning-Insurance.pdf

Insurers are using ML in automating the processes to increase operational efficiency from Claims registration to claims settlement. The claims processing time has been reduced and also, the Insurers can be more accurate on allocating the funds for claims settlement reserves. ML is also having application in giving insurance cover advice to customers depending upon their requirements.

5) Process Automation :
Process automation increases the operating efficiency by automating the repetitive tasks. Following are the automation applications in financial services.

I. Chatbots: Chatbots are the software applications that use artificial intelligence to mimic written or spoken human speeches via auditory or textual methods to simulate human interaction. Consumers of information in financial services range from customers to institutional officials. Once built correctly chatbots can perform 24/7 and hence no customer waiting for longer periods. Chatbots enable to understand the customer behavior in better way and thus help to design more accurate response system. Chatbots can provide meeting details, follow up details, links to documents and information sources, metrics and KPI through natural conversations. They can provide predictive insights across various sectors, right from Risk Management, Fraud prevention, regulatory reporting, sales insights to survey and market information, benchmarking and customer sentiment analysis. But Chatbots should be used as a complementary to customer experience.

II. Call center automation: Bots are useful in guiding the customers to serve themselves. They can be used as the first level support and answer the frequently asked questions. Also, ML and AI help to build the algorithms to learn, train and update these to predict customer response and provide relevant data to call center agent to solve complex queries.

III. Process Automation: Process automation eliminates the repetitive, cumbersome tasks done by the employees and thus increasing the efficiency and reducing costs. This frees the resources for more strategic and productive work. The newly developed Process automation tools analyze prior decisions and actions, learn from them and then take smarter decisions. One of the applications of process automation is event promotion where intelligent application tracks the location of the clients, predicts their chances of attending the event and sends them invitations.

Though machine learning has so many advantages over the traditional methods, many financial institutions are not effectively using the technology due to following reasons:

1) The ML and AI technology is having a high entry barrier due to expertise required and R&D in ML is costlier.

2) The required manpower is not available easily as lesser no. of people has required capabilities and skill set.
3) Updating existing infrastructure and retraining algorithms requires extra investment and efforts.
4) Before collecting the data, organizations should have clear idea about the expected outcome of the algorithms. Also, the quality of data used to train the algorithms is important. The data provided should be balanced and representative to avoid bias.
5) There is no universal ML algorithm which can be applied to all business cases and needs R&D.
6) The huge amount of data has created additional security risk of data leaks and security breach.

With the advancement in ML and AI the additional dependencies will be created due to the increased scalability of new technologies. This will in turn lead to emergence of more systematically important players in the market.

3. Conclusion

The rise in wealth of availability of data and its use along with development in ML and AI has made it possible for financial institutions to apply sophisticated techniques to serve the clients and customers. The aim of ML is not to replace Human resources entirely but to complement them to achieve maximum operational efficiency, reduce the costs and gain competitive advantage. The value mix of Traditional techniques, behavioral/sentiment analysis and ML techniques will give an edge to the institutions employing it. The major advantages of ML are decision making with prior learning experiences, avoidance of behavioral mistakes, and ability to handle large amounts of data in lesser time with high precision. The more efficient processing of information on credit risks, fraud detection and prevention, potential to increase regulatory compliance can give the financial stability to financial institutions. At the same time, the network advancements and scalability of new technologies may give rise to additional dependencies. Many current providers of ML and AI may fall outside the regulatory framework or may not be aware of the regulatory laws. The lack of interpretability of the ML and AI poses the risk of failure to predict and determine the potential effects and hence the algorithms can be trained to give optimal solutions in low volatility situations. There should not be use of opaque models which would result in unintended consequences. The scarcity of resources with required skills and knowledge is one more issue which would be faced by the institutions. The applications of ML and AI should be continuously monitored for their uses, innovations, regulatory and security implications.

4. References

1.Advisor, A. P. (n.d.). Machine Learning : A Revolution in Risk Management and Compliance ?

2. https://www.accenture.com/_acnmedia/PDF-84/Accenture-Machine-Leaning-Insurance.pdf

3. https://emerj.com/ai-podcast-interviews/machine-learning-for-credit-risk

4. https://www.valuewalk.com/2018/05/machine-learning-meets-portfolio-management