MACHINE LEARNING ALGORITHMS Towards financial services & analysis and fraud deduction

Dr.T.Lokeswara Rao
Professor & HOD of MBA
Swarnandhra College of Engineering & Technology
Narsapur, West Godavari District

Abstract:

Now a day's online transactions have become an important and necessary part of our lives. An occurrence of transactions is increasing; quantity of fraudulent transactions is also increasing rapidly. Financial services & analysis, in order to reduce fraudulent transactions, machine learning, like supervised learning and unsupervised learning etc., are discussed in this paper. The same set of algorithms are implemented and tested using an online dataset. Financial institutions also have been using ML tools for a number of applications.

Key words: Machine learning algorithms, financial services, financial analysis, fraud

Introduction

Machine Learning Algorithms (MLA) is quickly adopted for a variety of application in financial services. M L may be defined as a method of designing a progression of actions to solve the problem without human intervention, known as algorithms. The techniques can be used to find patterns in large amounts of data from innovative sources.

Cybercriminals are receiving smarter, and paradoxically, are leveraging advances in technology for their benefit. Financial institutions have no alternative, but to tighten their defenses and build up their capability earlier. They may well make additional complicated real-time insights possible on larger datasets, such as Internet-of-Things (IoT) sensors located around the world.

Definition

M L solutions are that they learn from training without being explicitly programmed. To put it simply, you need to choose the models and feed them with data. The model then automatically adjusts its parameters to get better outcomes.

In general, the more data you feed, the more correct are the result. Accidentally, huge datasets are very common in the financial services. There are peta bytes of data on transactions, customers, bills, money transfers, and so on. That is a fantastic fit for machine learning.

Types of Machine Learning

Machine learning types, they are: supervised unsupervised, deep and reinforcement learning.

Supervised learning is to start a relationship between two datasets and to use one dataset for estimate the other. Supervised learning algorithms are basically two varieties: regression and classification methods.

Regression method is a part of supervised learning method, this method to forecast outputs.

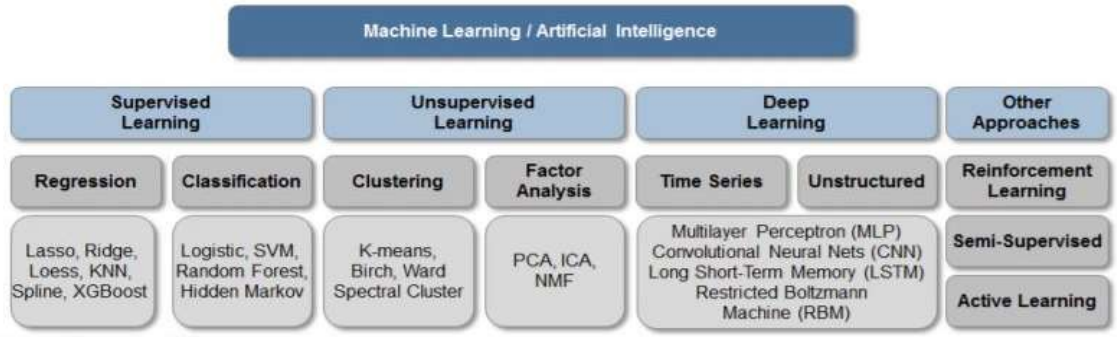
Classification method is the back and tries to recognize which category a set of classifications belong to.

Unsupervised learning is a machine learning technique to understand the structure of data and to identify the main drivers behind it. It will be used to identify the relationships between a huge numbers of variables, at a high level; these methods are categorize as clustering or factor analyses.

The goal of **clustering** is to reduce the amount of data by categorizing or grouping similar data items together.

Deep learning is to use multi-layered neural network to analyze a tendency. This systems will take on tasks that are hard for people to define but easy to perform. This learning is successfully an attempt to artificially recreate human intelligence. J.P.Morgan says deep learning is particularly well suited to the per-processing of unstructured big data sets. A deep learning model could use a hypothetical financial data series to estimate the probability of a market correction.

Reinforcement learning is an area of **Machine Learning**. It is about taking suitable action to maximize reward in a particular situation.


Machine Learning / Artificial Intelligence
Supervised Learning
Unsupervised Learning
Deep Learning
Other Approaches
Regression
Classification
Clustering
Factor Analysis
Time Series
Unstructured
Reinforcement Learning
Lasso, Ridge, Loess, KNN, Spline, XGBoost
Logistic, SVM, Random Forest, Hidden Markov
K-means, Birch, Ward Spectral Cluster
PCA, ICA, NMF
Multilayer Perceptron (MLP) Convolutional Neural Nets (CNN) Long Short-Term Memory (LSTM) Restricted Boltzmann Machine (RBM)
Semi-Supervised
Active Learning

Source: J.P.Morgan Macro QDS

Machine Learning Algorithms (MLAs)

Machine learning algorithms (MLAs) are programs that regulate themselves to perform better as they are showing to more data. The MLAs programs change how they process data over time, much as humans change how they process data by learning. So MLAs is a program with an exact way to adjust its own limitation, given feedback on its previous performance making prediction about a dataset.

Machine Learning in Finance

Machine learning in finance may work magic, even though there is no magic behind it. Still the success of machine learning project depends more on building efficient infrastructure, collecting appropriate datasets, and applying the exact algorithms.

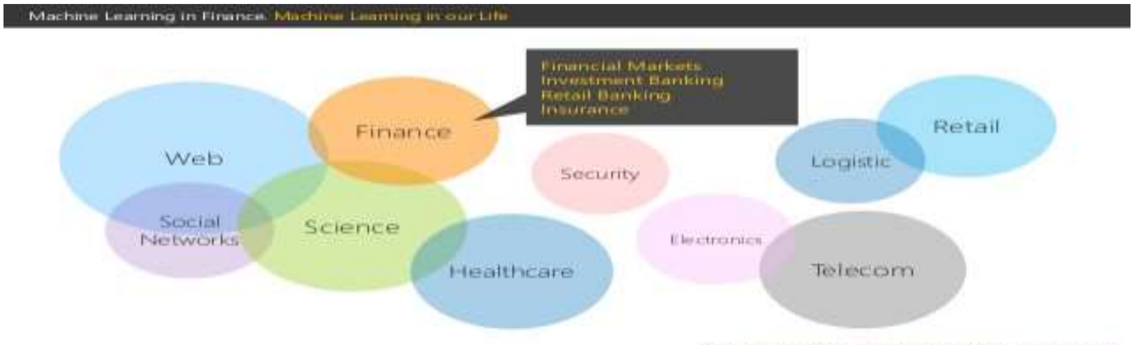
Financial institutions can use machine learning tools for a number of operational applications. Some of these applications include:

(i) Capital optimisation by banks,

(ii) Model risk management, and

(iii) Market impact analysis

The uses of machine learning should continue to be monitored. As the underlying technologies develop further, there is potential for more widespread use, beyond the use cases discussed in this report. It will be important to continue monitoring these innovations and to update this assessment in the future.



Proof: https://www.kaggle.com/wiki/DataScienceUseCases

The impact of fraud

Financial firms face a rapidly growing threat from cybercriminals – attacks persistently come in the form of money laundering, identity theft, and mobile fraud, among others.

One of the most common types of cybercrime, however, is bank and credit card fraud. The growth in e-commerce and mobile payments is partly behind the soaring incidence of card fraud in recent years, and the scale of the problem is vast. According to McKinsey, worldwide losses from card fraud could be close to $44 billion by 2025.

As well as the direct cost of fraud, companies are also suffering because of lost sales when genuine business are declined by fraud management systems. McKinsey suggests that false positives constitute up to 25 percent of declined transactions for e-commerce retailers.

The challenge for banks and financial institutions then is to quickly recognize and divide fake transactions from those that are legitimate, without impacting on customer experience.

Digital transactions

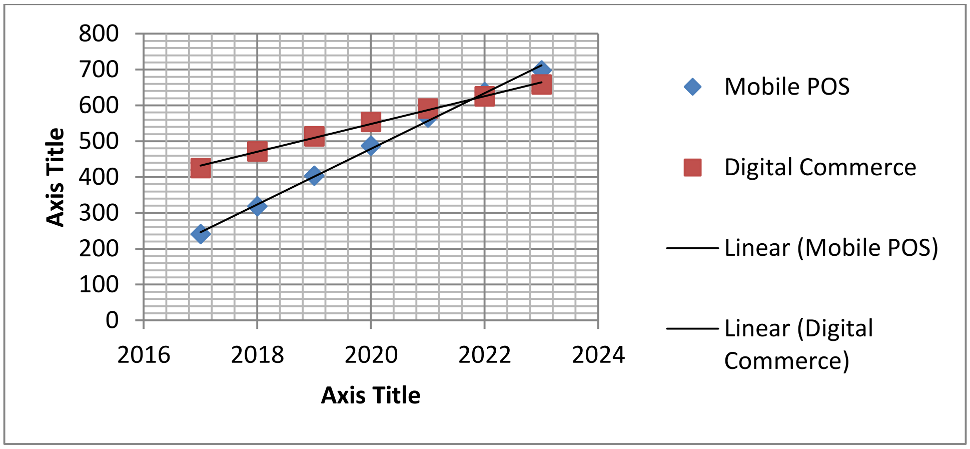
- Total transaction value in the digital payment segment amount to US$ 64,787M in 2019.
- Total transaction value is expected to show an annual growth rate (CAGR 2019-2023) of 20.1% resulting in the total amount of US$134,588M by 2023.
- The market's largest segment is digital commerce with a total transaction value of US$58,812M in 2019.
- From a global comparison perspective it is shown that the highest cumulated transaction value is reached in China (US$1,570,194M in 2019).

Digital transactions

- Total transaction value in the Digital Payments segment amounts to US$64,787m in 2019.
- Total transaction value is expected to show an annual growth rate (CAGR 2019-2023) of 20.1% resulting in the total amount of US$134,588m by 2023.
- The market's largest segment is Digital Commerce with a total transaction value of US$58,812m in 2019.
- From a global comparison perspective it is shown that the highest cumulated transaction value is reached in China (US$1,570,194m in 2019).

Mobile POS Payments Digital Commerce

Year	2017	2018	2019	2020	2021	2022	2023
Mobile POS	240.4	318.5	403.2	487.6	566.41	636.90	697.8
Digital Commerce	425.0	471.0	513.8	553.7	590.91	625.53	657.77



Reading Support In the Mobile POS Payments segment, the number of users is expected to amount to 697.8m by 2023.

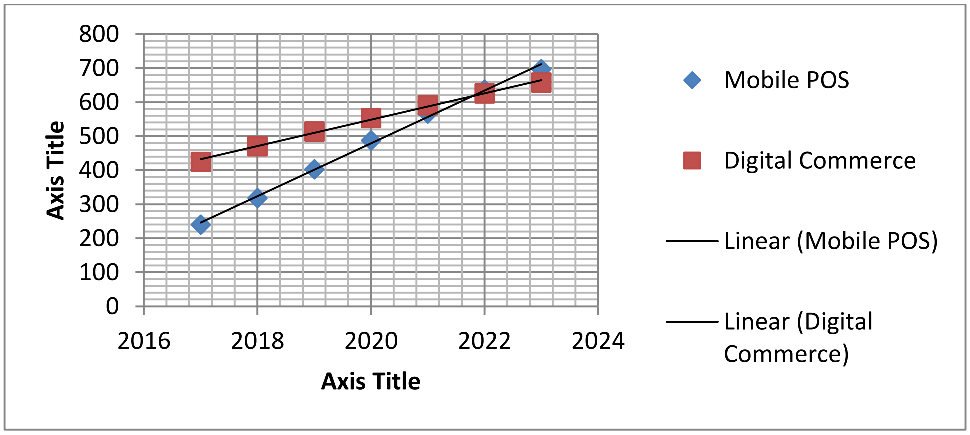
Details of credit and debit cards (in millions) 2017-19

Month & Year	Credit card (Millions)	Debit Card (Millions)	Total (Millions)
May 2019	48.9	824.9	873.8
Jun 2018	39.37	944	983.37
May 2017	30.86	880.03	910.89

Total number of ATMs in India 2,21,703, Cards users are increased, gradually transactions are increasing, quantity of fraudulent transactions is also increasing rapidly.

The Mobile POS Payments segment, the number of users is expected to amount to 697.8m by 2023.

Year	2017	2018	2019	2020	2021	2022	2023
Mobile POS	240.4	318.5	403.2	487.6	566.41	636.90	697.8
Digital Commerce	425.0	471.0	513.8	553.7	590.91	625.53	657.77

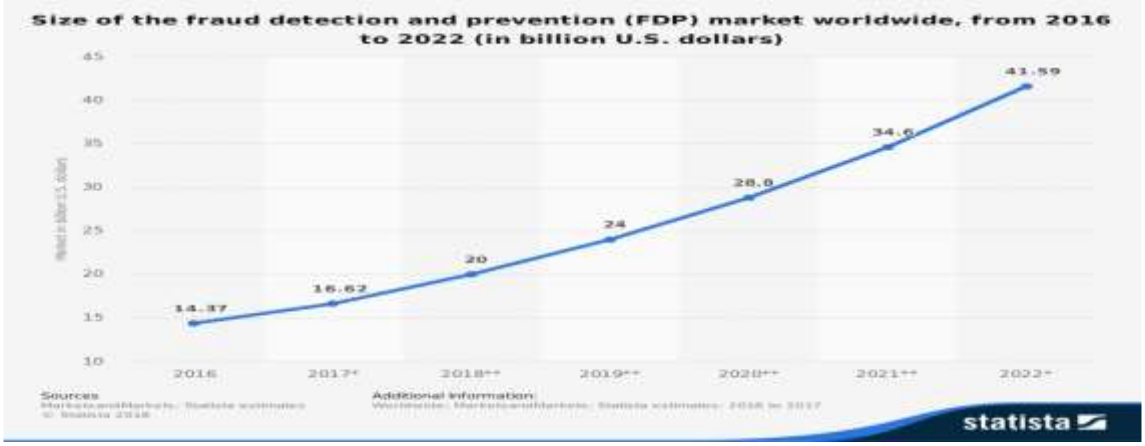


Fraud detection -Role of machine Learning

The financial services sector is on the eve of a major transformation, and the driving force in the rear it is (Artificial Intelligent) AI. Inventive applications for AI have already been found across area such as credit scoring, regulatory compliance, customer experience. Rapid advancements in technology, tasks that once took staff hours to complete manually can now be done in a matter of seconds.

Cybercriminals are getting smarter and, paradoxically, are leveraging advances in technology for their benefit. Financial institutions have no choice but to tighten their defenses and expand their own capabilities quicker.

According to Statista, in 2017, the global Fraud Detection and Prevention (FDP) market was estimated to be worth $16.6 billion. Areas where fraud detection and prevention are applied include insurance claims, money laundering, electronic payments, and bank transactions, both online and offline.



AI & ML, combined with the wealthy information available in the financial services sector, be providing organization with the means to keep their businesses, and defeat criminals.

Cluster analysis

Cluster analysis is part of the unsupervised learning. A cluster is a group of data that share similar features. The machine searches for comparison in the data. For example, if anyone can use cluster analysis for the following application:

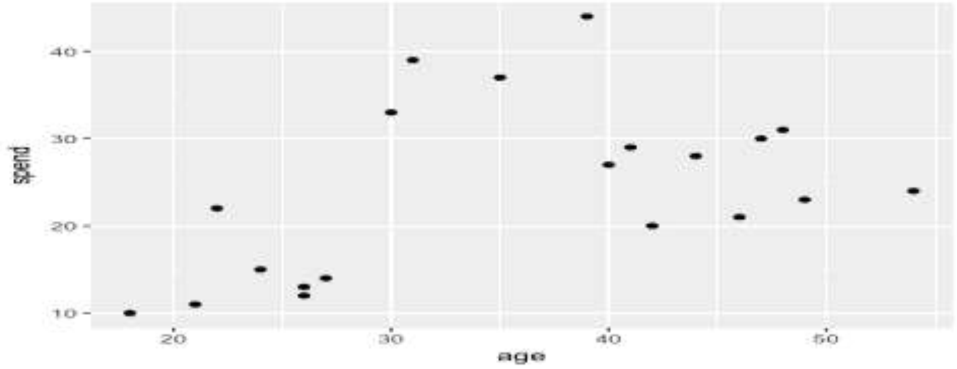
- Customer segmentation: Looks for similarity between groups of customers
- Stock Market clustering: Group stock based on performances
- Reduce dimensionality of a dataset by grouping observations with similar values

The most striking difference between supervised and unsupervised learning lies in the results. Unsupervised learning creates a new variable, the label, while supervised learning predicts an outcome. The machine helps the practitioner in the quest to label the data based on close relatedness. It is up to the analyst to make use of the groups and give a name to them.

Let's make an example to understand the concept of clustering. For simplicity, work in two dimensions. You have data on the total spend of customers and their ages. To improve advertising, the marketing team wants to send more targeted emails to their customers.

Library data

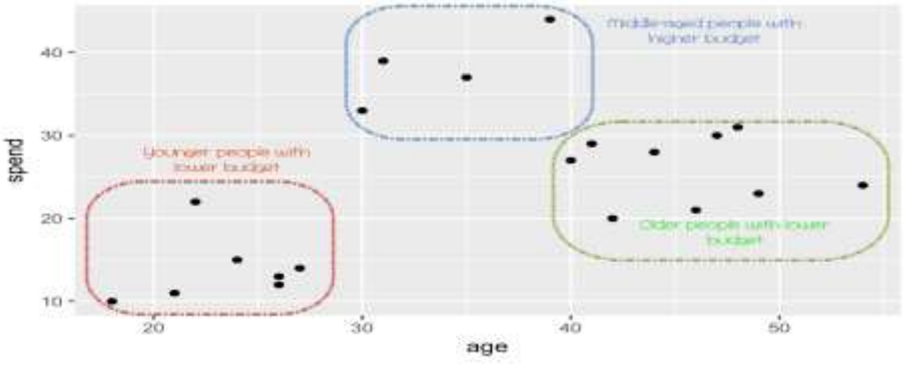
Age=X	18	21	22	24	26	26	27	30	31	35	39	40		42	44	46	47	48	49	54
Spend=Y	10	11	22	15	12	13	14	33	39	37	44	27		20	28	21	30	31	23	24



In the chart, you plot the total amount spend and the age of the customers

A pattern is visible at this point

1. At the bottom-left, you can see young people with a lesser purchasing power
2. Upper-middle reflects people with a job that they can afford spend more money
3. Lastly, elder people with a lower budget.



In the chart above, you cluster the observations by hand and define each of the three groups. This case is to some extent simple and very much visual. If new observations are appended to the data set, you can label them within the circles. You define the circle based on our decision. Instead, you can use ML to group the data objectively.

Conclusion

The apple of AI and ML technology is changing the provision of a number of financial services. Application of machine learning currently more widely used than other key FinTech innovations. In particular, financial services and financial analysis and fraud detection, emerge to be upward rapidly. Most market participant expect that AI and ML will be adopted further.

Machine Learning and AI, in general, are being adopted for a variety of applications in finance, excelling especially at fraud detection and stress testing. The ML algorithms banks now use, build upon familiar data science methods, such as linear regression, to handle millions of outputs, and utilize statistical methods to compress and summarize huge datasets. They show vast promise when their pitfalls, such as the lack of auditability, are accounted for and managed properly.

References

1. Heta Naik, “Credit card fraud detection for Online Banking transactions”, International Journal for Research in Applied Science & Engineering Technology, pp 4573- 4577, 2018

2. You Dai, Jin Yan, Xiaoxin Tang, Han Zhao and Minyi Guo, "Online Credit CardFraud Detection: A Hybrid Framework with Big Data Technologies", IEEE TrustCom/BigDataSE/ISPA , pp 1644 -1651, 2016

3. Suman Arora , "Selection of Optimal Credit Card Fraud Detection Models Using a Coefficient Sum Approach" , International Conference on Computing, Communication and Automation (ICCCA2017), pp 482 - 487, 2017.

4. Nobuchika Mori (2017), “Will FinTech create shared values?” speech at Columbia Business School conference, May.

5. Defined in EIOPA (2017), “Opinion of the Occupational Pensions Stakeholder Group on JC Big Data,” EIOPA-OPSG-17- 06 15, March, pp. 6-7. See also U.S. Federal Trade Commission (2016), “Big Data: A Tool for Inclusion or Exclusion,” January, p. 3.

6. See EIOPA (2017); U.S. Federal Register (2017), Vol. 82, No.33, and Bureau of Consumer Financial Protection: Docket No. CFPB Notice and Request for Information Regarding Use of Alternative Data and Modelling Techniques in the Credit Process, February 21, 2017 (“CFPB RFI”); European Banking Authority (2017), “Report on innovative uses of consumer data by financial institutions, June. See also FSB FinTech Issues Group (2017), p. 19.

7. OECD (2013), “Guidelines on the Protection of Privacy and Transborder Flows of Personal Data,” July.