IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Identifying and exploring areas in the Retail Industry that could apply machine learning Algorithms and suggesting algorithms

Sigma Nasit

Abstract

The retail industry faces many challenges such as : right product fit, small basket value, poor management of inventory, right recommendation systems etc, all these things are connected by a common thread i.e it is all human thinking. Thus these become a perfect fit for Machine Learning (a subset of Artificial Intelligence) that can mimic human thinking and thus provide predictions that can be used by businesses to optimize themselves. In this paper we will be discussing the common problems faced in the retail industry and in what area.We will be exploring how and which Machine Learning algorithms can be used for a specific problem, which will be the best fit ,some variables in performance with different algorithms and feasibility of these algorithms.

Keywords : Machine Learning , Data, Retail Industry, Marketing, Inventory Management, Pricing, Data Analytics, Artificial Intelligence

Introduction

Understanding Machine Learning :

The scientific discipline of machine learning enables computers to learn without explicit programming. The ability to learn is what, as the name suggests, gives the computer a more human-like quality. Today, machine learning is being actively used, possibly in a lot more places than one might think.Types : Supervised, Unsupervised, Reinforcement.

However, unlike humans, it needs a lot of data to learn. Computers can use machine learning to detect patterns in the data that are both obvious and hidden after being taught over such enormous amounts of

data. This is the reason why machine learning is quickly being adopted by many huge organizations, since it can quickly and efficiently comb through large amounts of data to identify hidden new patterns more rapidly than a human although mimicking human thinking.

Exploring Areas :

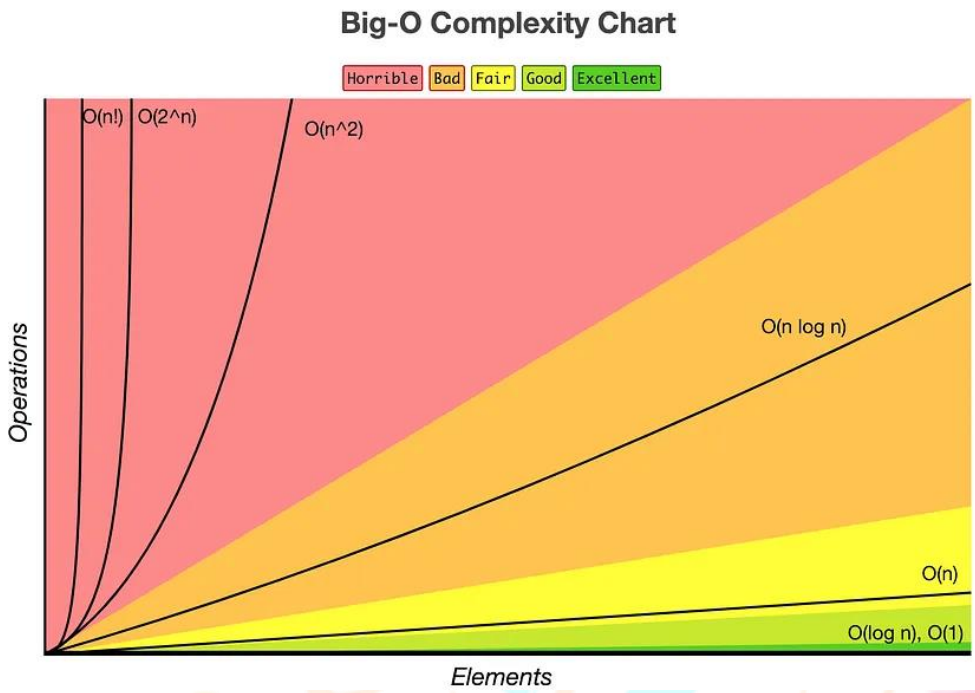
Listing and understanding requirements of Machine Learning :

- **Data :** Data is on of the key component for Machine Learning. It is element on which the algorithms will be run and results in form of patterns will be obtained. There are two types of data :

    1. Labeled Data : This consists of as the name suggests labeled data, i.e the data has meaningful tags that makes not only the data preprocessing easier but also helps in faster implementation of certain algorithms. Algorithms(Supervised) used are generally Regression and Classification.This type of data although required in less quantity as compared to unlabeled data, preparation of the same is difficult and time consuming.

    2. Unlabeled Data : This consists of as the name suggests data that is unlabeled.It does not contain any meaningful tags. Here the task for the algorithms lengthens a bit as it needs to sort or group or separate data in a manner that makes it useful for the algorithm to gain insight.Algorithms(Unsupervised) used are generally Clustering, Association, Dimensionality Reduction. This type of data is required in large quantities as useful pattern recognition comes over large data. Although could be time consuming it helps in find hidden patterns and trends.

- **Processing Time :** It indicates the time an algorithm takes to process data. Referred to as time complexity in computer science terms. Time complexity can be thought of as a gauge for how quickly or slowly an algorithm will operate given a given input size. Time complexity is always expressed in terms of an input size, let's say n.Using the Big O Notation, which specifies an upper bound of an algorithm that bounds a function solely from above.Refer fig(1).



fig(1)

- **Processing Space :** It refers to the amount of space that is required by the Machine Learning Algorithm for processing. Referred to as Space Complexity in Computer Science terms.You might think of space complexity as the additional memory needed to run your method. Similar to time complexity, it is specified in relation to an input size (n).Notation used : Big ORefer to fig(1) for more information.

- **Computational Power :** Here referred to as the computational power required to execute the algorithm. Many algorithms are a costly affair in terms of parallel and constraint execution that requires to set up additional computational units apart from the primary unit.

Methods :

For every area explored, the problem is explored, the proposed application of machine learning is given, an algorithm is chosen based on the ratings and briefly explained along with it working.The primary that would be necessary to be fed in the algorithm for business goals would be mentioned. Reasons to choose the particular algorithm will be mentioned and a little comparison in terms of how majorly the chosen algorithm is better than the other widely used algorithms will be mentioned.

For every area explored a table is prepared with the parameters of requirements of Machine Learning algorithm i.e data, processing time, processing space and computational power. Apart from these parameters three additional parameters : cost (economic viability) , ease of use (operational viability) and additional functionality (any other additional functionality that the algorithm provides) will be considered for rating the algorithms.

These algorithms will be rated theoretically. The ratings will be from 1-5 , 1 being the lowest and 5 being the highest i.e for example the operational cost needs to be ideally low for it to be desirable and thus low

operational cost would result in high ratings,i.e the ratings will vary in accordance to the parameter's desirable result.

Exploring Areas :

In this section we will be seeing some of the areas of business that require solutions. We will also match it with above mentioned Machine Learning requirements and proposing a Machine Learning Algorithm for the same :

1. Demand Prediction

Problem exploration : Retailers (especially MSMEs) face a huge loss due to overstock or understock of products leading to either loss/money invested or failing a potentially profitable market respectively.

Proposed Application: Machine Learning can be applied to the above problem statement as the Machine Learning Algorithm can process data and identify trends in consumption of products. The business authority can feed into the algorithm customer purchasing data and as a result obtain trends in purchasing that then can be used to manage stock, prepare for future markets and cater to untapped customer needs.

Algorithm : LSTM (Long Short Term Memory) : It is a type of Recurrent Neural Network(RNN) that specializes in dealing with sequential or time based data as it contains a memory cell that can hold information for extended periods of time which can then be used as long term dependencies for future prediction.

Working : Consider the current input along with the prior concealed state and internal cell state.
The four separate gates' values can be determined by doing the following calculations: For each gate, multiply the corresponding vector with its corresponding weights element-by-element to determine the parameterized vectors for the current input and the prior hidden state.
Apply the appropriate activation function to the parameterized vectors for each gate element. The list of gates with the activation function to be used for the gate is provided below.
Calculate the element-wise multiplication vectors of the input gate and the input modulation gate, followed by the element-wise multiplication vectors of the forget gate and the previous internal cell state to determine the current internal cell state.

Data to be used : Labeled data of products with their purchasing history and corresponding prices at that point of time.

Reason to choose this algorithm :

1. Long term dependencies can be captured effectively using this algorithm which becomes essential in this case as we need to analyze over sequential time series data .So that the business authority does not need to go for a complex algorithm with working deficiencies.

2. As it is a neural network it can efficiently form useful connections based on target function which makes it easier to identify trends over a period of time.Thus the business authority need not constantly monitor it ,saving time.
3. The memory is present in an organized network thus decreasing access time.
4. As the algorithm consists of memory cells we can capture data over a long time and calculate its dependency to current scenario, including essential past information.Thus the business authority as per needs can include or remove time periods.

Algorithm	Data viability	Processing Time	Processing Space	Processing Power	Economic viability	Operational viability	Additional Functionality	Total (out of 35)
LSTM	4	5	5	3	3	4	3	**27**
ARIMA	4	3	3	3	4	3	3	**23**
ES	4	3	4	4	4	2	1	**22**
STD	4	3	4	3	4	2	1	**21**

Fig(2) : theoretical comparative rating for demand prediction algorithms

Compared with other widely used algorithm :

1. Performs better than ARIMA (Auto-Regressive Moving Average) as it efficiently deals with gaps in data.
2. LSTM overthrows Exponential Smoothing(ES) as it lags, i.e it is not dynamic in nature.
3. Although in some cases Seasonal Trend Decomposition(STD) may prove to be good, for the feature that the time cycle is to be input by the business authority , it can be a repetitive task as and when the time period increases making it a tedious task.

2. Dynamic Pricing :

Problem Exploration : Retailers (both online and offline) many times fail to keep up with the market trends and even when they manage to do so they fail to give their customer a reasonable (yet profitable for themselves) price, resulting in lack of sale or loss.

Proposed Application : Machine learning algorithms can be applied in this particular situation by letting the algorithm explore over huge datasets of a product in the market, it's availability and it's demand based on which the algorithm outputs a reasonable price, which is market robust as well as profitable as well as fitting for the customer.

Algorithm : Reinforcement learning(RL) : It involves acting appropriately to maximize reward in a certain circumstance. It is used by a variety of programmes and computers to determine the optimal course of action to pursue in a given circumstance. There is no correct answer in reinforcement learning, but the reinforcement agent decides what to do to complete the task. This is different from supervised learning, where the training data includes the answer key and the model is trained with that answer. It is obligated to gain knowledge from its experience in the absence of a training dataset.

Reinforcement Learning (RL) is the science of decision making. It involves understanding how to act in a situation to reap the most benefits. Data for RL is gathered from machine learning systems that employ a trial-and-error process.

Algorithms used in reinforcement learning determine the next course of action based on results. The algorithm receives feedback after each step that aids in determining whether the decision it made was good, bad, or indifferent. It is a useful method for automated systems that must make numerous tiny judgments without human supervision.

Working : The interaction between a learning agent and its environment is defined by a formal framework in reinforcement learning in terms of states, actions, and rewards. This framework aims to offer a straightforward method of expressing key aspects of the artificial intelligence challenge.

The following are components of reinforcement learning: policy, reward function, value function, model of the environment.

Policy: Policy outlines the behavior of learning agents for a specific time frame. It is a mapping between perceived environmental states and the activities that should be conducted in such states.

Goal definition in a reinforcement learning problem is done using the reward function.A reward function is a function that calculates a score based on how the environment is doing.

Value functions define what is beneficial over the long run. The total quantity of rewards an agent might anticipate accumulating over the course of their existence, starting from a certain state, is the value of that state.

a simulation of the environment Plans are made using models.

Basically upon performing a task the and based on the result predicted and upon receiving feedback it improves itself and adapts to the changing environment.

Data to be used : Consumption of the product over a time period, import/export of the product, demand of the product, seasonality .

Reasons to choose this algorithm :

1. We can efficiently use reinforcement learning algorithm to assist in preparing the dataset for the problem statement as well as it can take into consideration different parameters for similar products such as products providing with the same functionality but with improved quality or different raw material or different processing technique .
2. Our data would be inconsistent with time over a large period of time, and may also contain products with shorter market periods or similar products or products being sold by different means (online , offline, wholesale etc).

3. As the algorithm is such that it uses trial and error with a reward based system the result and conclusion obtained from it becomes more fluid in nature ,i.e it is not rigidly constrained by the algorithmic parameters.

4. This algorithm is known for its decision making abilities and how well it can handle non deterministic situations, i.e situations that are not encountered during training of the algorithm, thus making it easier for it to work during uncertain market conditions as well.

Algorithm	Data viability	Processing Time	Processing Space	Processing Power	Economic viability	Operational viability	Additional Functionality	Total (out of 35)
RL	4	4	4	4	4	5	4	**29**
BM	3	3	2	3	4	3	2	**20**
DT	3	1	2	3	2	2	3	**16**

Fig(3) : theoretical comparative rating for dynamic pricing algorithms

Compared with other widely used algorithm :

1. Bayesian Model(BM) requires to input a base value from the business authority(making it a new set of task to analyze and have a optimal base value, such that it does not interfere with the pricing and moves over or under this predefined base value based upon past prices and thus neglects the current conditions and circumstances, thus making it less/rigidly dynamic in nature, whereas on the contrary the reinforcement learning model is dynamic in nature taking in account the current as well as changing environment (market).

2. Decision Tree algorithm(DT) becomes a bit more constrained when it comes to the parameters to be considered for prediction, it requires the business authority to decide on the parameters and works strictly around them, which can be good if we consider niche markets but for large markets reinforcement learning model not only decides on the parameters innately but also finds hidden factors that may affect the outcome**.**

3.Recommendation Engine

Problem Exploration : Online retailers although enjoy high traffic often fail in terms of presenting the customer with the right product fit, resulting in high traffic yet less conversion to sales incurring loss as well as high cost to conversion ratios.

Proposed Solution : Machine learning algorithms can design Recommendation Engines that can help solve this matter, where it by tracking the users recent and past activities as well as similar user consumption patterns would help recommend the user a product more suitable to their particular taste, which leads to higher customer retention as well as conversions to sale.

Algorithm : BIRCH (balanced iterative reducing and clustering using hierarchies) : It accomplishes hierarchical clustering over big data sets. The expectation-maximization approach can be modified to speed up the k-means clustering and Gaussian mixture modeling processes. BIRCH has the advantage of being able to dynamically and progressively cluster incoming multi-dimensional metric data points to provide the best clustering possible given the available resources (memory and processing time). BIRCH typically only needs to scan the database once.

Working : This algorithm uses the CF (clustering features) tree as a foundation. Additionally, this approach builds clusters using a tree-structured summary.The algorithm compresses the data into sets of CF nodes in relation to the CF tree. CF subclusters are nodes that contain several sub-clusters. These CF subclusters are located in CF nodes without terminals.

The CF tree, a height-balanced tree that manages clustering characteristics and stores the relevant data for additional hierarchical clustering, collects and manages clustering features. This eliminates the requirement to work with the entire set of input data. Three numbers (N, LS, and SS) are used to represent the tree cluster of data points as CF.

Data to be used : Customer purchase data

Reasons to choose this algorithm :

1. It uses clustering to form clusters/ classes of product thus making the data set more viable to work on and extract information for future prediction.
2. It uses clustering at its core to group data points ,thus cutting down on the performance cost for separating new data points.
3. The data is stored in form of little summaries thus allowing incorporation of multiple parameters and thus more detailed connection and correspondence between a customer and their selection of products.
4. It organizes data in a hierarchical manner, meaning that not all parameters have the same weightage in the selection of a particular product,i.e the algorithm considers the weights of the parameters during calculations, thus preventing faulty or erroneous predictions due to over/under consideration of role of a parameter in purchase of the product.

Algorithm	Data viability	Processing Time	Processing Space	Processing Power	Economic viability	Operational viability	Additional Functionality	Total (out of 35)
BIRCH	4	4	4	4	4	4	3	**27**
CL	3	4	2	3	3	3	2	**20**
ANN	4	4	3	3	3	3	4	**24**

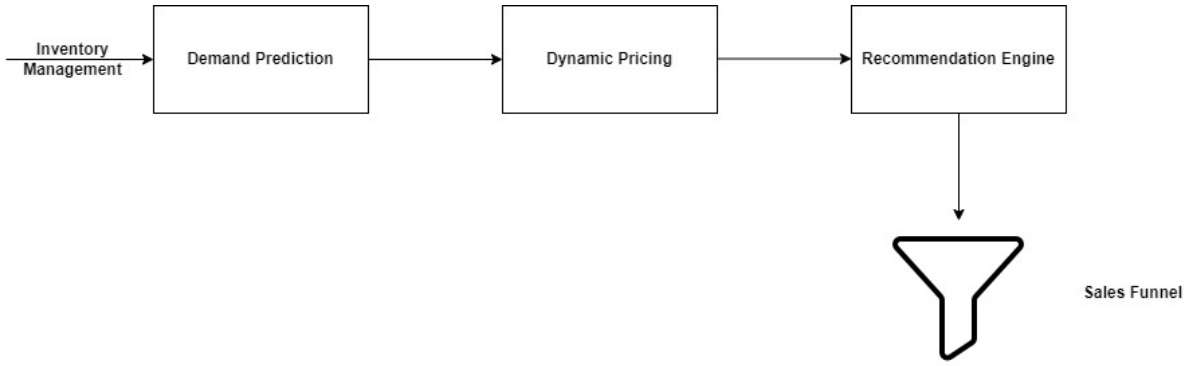
Fig(4) : theoretical comparative rating for recommendation system algorithms

Compared with other widely used algorithm :

1. Clustering(CL- referred to as in fig(4)) of any type is generally used for this purpose but it is very constrained and one dimensional (focussing on singular parameter) in nature, whereas BIRCH incorporates multiple parameters along with their actual weightage.

2. Deep Learning Algorithms(Artificial Neural Networks - ANN) are although know for making connections also demands space as well as computational power whereas in BIRCH algorithm somewhat same results can be obtained with much less operational costs.

Future Scope : The areas explored above are a part of the business pipeline and can be worked as a framework in the following manner :



Inventory management can be assisted by demand prediction, i.e the business authority can manage their inventory based on the demands of the current markets as well as prepare for future markets.Moving on , based on the inventory as well as the demand analysis of the market the business can have reasonable prices for the customer which prove to be profitable for the business at the same time, all the while ensuring stock of the product as well as calculating the time till which stock can be kept and what time and what price it should be sold using dynamic pricing. To make it complete recommendation engines can be used to boost sales.

Citations :

1. https://research.aimultiple.com/dynamic-pricing-algorithm/
2. https://www.itransition.com/machine-learning/retail
3. https://itnext.io/what-are-the-top-recommendation-engine-algorithms-used-nowadays-646f588ce639
4. Role of input features in developing data-driven models for building thermal demand forecast :Citation DataEnergy and Buildings, ISSN: 0378-7788, Vol: 277, Page: 112593,Publication Year2022
5. DEMAND FORECASTING: AI-BASED, STATISTICAL AND HYBRID MODELS VS PRACTICEBASED MODELS - THE CASE OF SMES AND LARGE ENTERPRISES : **Ternopil** Vol. 15, Iss. 4, (2022): 39-62.
6. Mario Angos Mediavilla, Fabian Dietrich, Daniel Palm,
7. Review and analysis of artificial intelligence methods for demand forecasting in supply chain management,Procedia CIRP,Volume 107,2022,Pages 1126-1131,ISSN 2212-8271,https://www.sciencedirect.com/science/article/pii/S2212827122004036?ref=pdf_download&fr=RR-2&rr=80ff4d46f9bd85ad#section-cited-by
8. Mario Angos Mediavilla, Fabian Dietrich, Daniel Palm,Review and analysis of artificial intelligence methods for demand forecasting in supply chain management,Procedia CIRP,Volume 107,2022,Pages 1126-1131,ISSN 2212-8271,https://www.sciencedirect.com/science/article/pii/S1877050922003076?ref=pdf_download&fr=RR-2&rr=80ff503d39d685ad

9. Mario Angos Mediavilla, Fabian Dietrich, Daniel Palm,Review and analysis of artificial intelligence methods for demand forecasting in supply chain management,Procedia CIRP,Volume 107,2022,Pages 1126-1131,ISSN 2212-8271,
10. R. Abbate, P. Manco, M. Caterino, M. Fera, R. Macchiaroli,Demand forecasting for delivery platforms by using neural network,IFAC-PapersOnLine,Volume 55, Issue 10,2022,Pages 607-612,ISSN 2405-8963,https://www.sciencedirect.com/science/article/pii/S240589632201758X?ref=pdf_download&fr=RR-2&rr=80ff5485a91c85ad
11. http://idm-lab.org/bib/abstracts/papers/aaai93.pdf
12. https://par.nsf.gov/servlets/purl/10186768
13. Arnoud V. den Boer,Dynamic pricing and learning: Historical origins, current research, and new directions,Surveys in Operations Research and Management Science,Volume 20, Issue 1,2015,Pages 1-18,ISSN 1876-7354,https://www.sciencedirect.com/science/article/abs/pii/S1876735415000021
14. Ahmad Faruqui, Ryan Hledik, John Tsoukalis,The Power of Dynamic Pricing,The Electricity Journal,Volume 22, Issue 3,2009,Pages 42-56,ISSN 1040-6190,https://www.sciencedirect.com/science/article/abs/pii/S1040619009000414
15. P.K. Kannan, Praveen K. Kopalle (2001) Dynamic Pricing on the Internet: Importance and Implications for Consumer Behavior, International Journal of Electronic Commerce, 5:3, 63-83, DOI: 10.1080/10864415.2001.11044211
16. F.O. Isinkaye, Y.O. Folajimi, B.A. Ojokoh,
17. Recommendation systems: Principles, methods and evaluation,Egyptian Informatics Journal,Volume 16, Issue 3,2015,Pages 261-273,ISSN 1110-8665,https://www.sciencedirect.com/science/article/pii/S1110866515000341
18. JOUR, Kumar, Pushpendra,Thakur, Ramjeevan Singh.,2018, 2018/12/01, Recommendation system techniques and related issues: a survey, International Journal of Information Technologyhttps://link.springer.com/article/10.1007/s41870-018-0138-8#citeas
19. Yong Yu, Xiaosheng Si, Changhua Hu, Jianxun Zhang; A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures. *Neural Comput* 2019; 31 (7): 1235–1270. doi: https://doi.org/10.1162/neco_a_01199
20. Hossein Abbasimehr, Mostafa Shabani, Mohsen Yousefi,An optimized model using LSTM network for demand forecasting,Computers & Industrial Engineering,Volume 143,2020,106435,ISSN 0360-8352,https://www.sciencedirect.com/science/article/abs/pii/S0360835220301698
21. An ARIMA Supply Chain Model,Kenneth Gilbert,**Published Online:**1 Feb 2005https://doi.org/10.1287/mnsc.1040.0308
22. Forecasting using simple exponential smoothing method,Ostertagová, Eva; Ostertag, Oskar. Acta Electrotechnica et Informatica**; Kosice** Vol. 12, Iss. 3, (2012): 62. DOI:10.2478/v10198-012-0034-2
23. http://www.nniiem.ru/file/news/2016/stl-statistical-model.pdf
24. Bayesian strategies for dynamic pricing in e-commerce,Eric Cope,First published: 22 December 2006 https://doi.org/10.1002/nav.20204
25. Sandeep Saharan, Seema Bawa, Neeraj Kumar,Dynamic pricing techniques for Intelligent Transportation System in smart cities: A systematic review,Computer Communications,Volume 150,2020,Pages 603-625,ISSN 0140-3664,https://www.sciencedirect.com/science/article/abs/pii/S0140366419310990
26. Zhang, T., Ramakrishnan, R. & Livny, M. BIRCH: A New Data Clustering Algorithm and Its Applications. *Data Mining and Knowledge Discovery* 1, 141–182 (1997). https://doi.org/10.1023/A:1009783824328
27. Putri, D.C.G., Leu, JS. (2020). Towards the Implementation of Movie Recommender System by Using Unsupervised Machine Learning Schemes. In: Deng, DJ., Pang, AC., Lin, CC. (eds) Wireless Internet. WiCON 2019. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 317. Springer, Cham. https://doi.org/10.1007/978-3-030-52988-8_9