

MACHINE LEARNING APPROACH - A SCIENCE TO MAKE SYSTEM SMART- LITERATURE REVIEW

MRS. DIPTI NASHINE

Assistant Professor ,DYPIMCA, Akurdi,
Email id:Dipti.nashine@gmail.com

MRS. KRUPALIDHAWALE

Assistant Professor ,DYPIMCA, Akurdi,
Email id:krupali.acem@gmail.com

ABSTRACT:

One of the widely used research area in today's world is artificial intelligence and one of the scope full area is Machine Learning (ML). This is literature review paper describing ML, the process of machine learning. It is a science making system (computers) to understand from the past behaviour of data or from historic data to behave smartly in every situation like human beings do. Without having the same type o situation a human can behave or can react to the condition smartly. So with the help of machine learning also an effort is been taken to make and act computer smart getting computers to learn and act like humans do, by serving them data and information in the form of observations and real-world communications. This paper also gives the some application areas where ML is already working successfully.

KEYWORD: Machine learning, machine learning process

INTRODUCTION

Machine Learning is the latest way to make our system, computer, home appliances and other devices smart. SMART system are the system or devices which are used to make best use of knowledge and experience and inculcate all these past experience to make a machine react to the situation in the best possible manner a human being can. It is the used to optimize the resources, give results with more accuracy.

Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning algorithm makes efforts on the development of computer application and programs which can access data and can be use to learn for themselves[17].

Machine Learning Basic Concepts

Machine learning is a branch of artificial intelligence that employs a variety of statistical, probabilistic and optimization techniques that allows computers to "learn" from past examples and to detect hard-to-discern patterns from large, noisy or complex data sets. This capability is particularly well-suited to medical applications, especially those that depend on complex proteomic and genomic measurements [19].

There are many different types of machine learning algorithms, with hundreds published each day, and they're typically grouped by either learning style (i.e. supervised learning, unsupervised learning, semi-supervised learning) or by similarity in form or function (i.e. classification, regression, decision tree, clustering, deep learning, etc.).

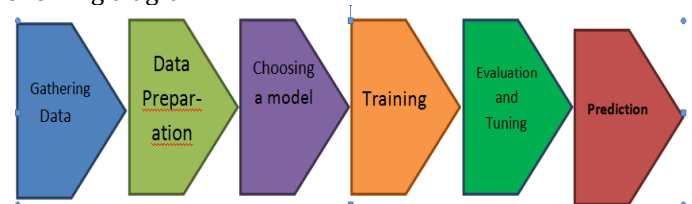
Regardless of learning style or function, all combinations of machine learning algorithms consist of the following [18]:

- Representation (a set of classifiers or the language that a computer understands)
- Evaluation (aka objective/scoring function)
- Optimization (search method; often the highest-scoring classifier, for example; there are both off-the-shelf and custom optimization methods used)

The fundamental goal of machine learning algorithms is to generalize beyond the training samples i.e. successfully interpret data that it has never 'seen' before [18]. Machines have allowed us to do complex computations in short amounts of time. This has given rise to an entirely different area of research which was not being explored: teaching machines to predict a likely outcome by looking at patterns. Machine Learning has many application areas and is being used to solve problems almost ranging from predictions of result at NSE-BSE, stock market, medical field to rainfall prediction. There are multiple famous machine learning algorithms in use today and new algorithms are popping up every other day.

Machine Learning Process

The machine learning process is considered to be bit tricky and challenging as well. We rarely find the machine learning process easy, but the process which is widely followed for machine learning is given in the following diagram.



1. **Gathering Data:** We consider the data gathering step to be the basic step for machine learning process. Though it is a preliminary step but it is very important because the quality and quantity of data that we are expecting out of this step will directly help in determining the predictive model.

2. **Data preparation:** After the data is gathered from the source, next step is the stacking data takes place and data is prepared so that, now our data can be use in machine learning training process. Relevant visualization of data can be done as it is good time to do and find if there are any relevant relationships between different variables or any data imbalances.

Data which has been gathered so far is now to be splitted into two parts. The first part, used in training our model, will be the majority of the dataset. The second part will be used for evaluating our trained model's performance. We don't want to use the same data that the model was trained on for evaluation, since it could then just memorize the "questions", just as you wouldn't use the same questions from your math homework on the exam. Sometimes the data we collect needs other forms of adjusting and manipulation. Things like de-duping, normalization, error correction, and more. These would all happen at the data preparation step.

3. **Choosing a model (Selecting a model):** Next step in the machine learning process is to select a model relevant to the research study. Many models have been created by researchers and data scientists over the years. From these model some are very appropriate for working with image data, others found its relevancy in sequences (such as text, or music), some are suitable for numerical data, others found relevancy for text-based data.

4. **Training:** Now one of major step in the process of machine learning is —the Training. This is the step which, we will use for our data to cumulatively in progression to improve the ability of our model to predict.

This can be related to a person learning to drive a vehicle. In the beginning the learner does not know how the parts like accelerator break and clutch works, or when they should be used or how to coordinate among these three. However, after committing lots of mistakes, rigorous practice and regular correction a licensed driver emerges. Moreover, after a year of driving, they've become quite adept. The performance of driving and responding to real-world data has changed their driving capabilities, improving their skills.

5. **Evaluation:** After the training process is successfully completed now it's time to evaluate the model to check its accuracy. At this moment the dataset that we had kept aside comes into action. By evaluation process, we are allowed to test user defined (our) model against data that has never been used for training. This metric allows us to see how the model might perform against data that it has not yet seen. This is meant to be representative of how the model might perform in the real world.

Parameter Tuning: Once you've done evaluation, it's possible that you want to see if you can further improve your training in any way. We can do this by **tuning our parameters**. There were a few parameters we implicitly assumed when we did our training, and now is a good time to go back and test those assumptions and try other values.

6. **Prediction:** Last step of Machine learning process is using data to answer the questions So **Prediction**, or inference, is the step where we get to answer some questions. This is the point of all this work, where the value of machine learning is realized.

Machine Learning process may differ slightly as per the different condition and the system selected for the

prediction. But the basic flow of the process is always followed.

Application area of Machine Learning:

There are many areas in which machine learning is successfully implemented

1. **Diagnosis in Medical domain:** Machine Learning provides variety of methods, tools and techniques that can help solving diagnostic and prognostic problems in a variety of medical domains. It is being used for the analysis of the importance of clinical parameters and of their combinations for prognosis, e.g. prediction of various diseases and its progression, for the extraction of medical knowledge at deepest level so that it can be used for outcomes research, for planning therapy and support, and for overall patient disease management. ML is also used for data analysis, like detection of regularities in the data by appropriately dealing with imperfect data, interpretation of continuous data used in the Intensive Care Unit(**datasciencecentral, 2017**), and for planting smart alarm system which can result in effective and efficient monitoring.

To improve the efficiency, effectiveness and the quality of medical care in form of medical reports, Machine Learning methods can be used to help computer-based software and systems in the healthcare department and area by providing opportunities to facilitate and enhance the work of medical experts and ultimately.

In medical examination the main focus is in finding the occurrence of a disease followed by its correct identification. There is a different category for each (one) disease under consideration and one category for cases where no disease is present. Machine learning improves the correctness of medical finding by analyzing records of a diseased person.

The dimensions here are typically the results of some medical examinations (such as BP monitoring results, temperature and all sort of blood tests) or medical diagnostics (for example medical imageries), presence/absence/intensity of various symptoms and basic physical information about the patient (age, sex, weight etc). On the basis of the results of these measurements, the doctors narrow down on the disease inflicting the patient.

Machine learning is frequently used in cancer diagnosis and detection. Recently machine learning application has been seen in cancer prognosis and prediction. This latter approach is particularly interesting as it is part of a growing trend towards personalized, predictive medicine. In assembling this review we conducted a broad survey of the different types of machine learning methods being used, the types of data being integrated and the performance of these methods in cancer prediction and prognosis.[1].

2. **Identification of Image:** Another application in the field of machine learning is identification of images. Many conditions are their where one has to identify among the set of images.

- Detection of appearance (face) – This is one of the most challenging job in machine learning. As there can be two different aspects, either to find a existence of face or no face, or to find the exact appearance from the huge set

of database. A face detection case may have a huge database of different faces of different people.

- Identification of character in text – Next segment in identification could be character recognition. Small images could be consider here, each comprising a one character. The groupings can consist any of ten digits, 26 letters of alphabet along with special letterings

3. Speech Acknowledgment: It is the conversion of vocal words into text. It is also known as “automatic speech recognition” (ASR), “computer speech recognition” Text to Speech (TTS) or “speech to text” (STT)

In this application, software can recognize spoken words. The measurements in this application might be a set of numbers that represent the speech signal. We can segment the signal into portions that contain distinct words or phonemes. In each segment, we can represent the speech signal by the intensities or energy in different time-frequency bands (**datasciencecentral, 2017**).

4. Automated trading strategy: In finance, it refers to small term exchange and it involves great measures of cares. User attempts to implement a trading algorithm considering lot of securities on the basis of magnitudes such as historical data (collected from past transaction) and general economic variables. These measurements can be cast as a classification or estimation problem. Here the hypothesis regarding the price is that it will move on the way to a past average (**datasciencecentral, 2017**)

5. Learning by establishing Associations: It is a process which helps in establishing an association between the different products or habits and learning those habits to predict something useful. This is widely used application where some unmatched or not related may show some relation or connectivity to one another. Once the pattern of buying is analysed

One of the widely used applications of machine learning is studying the association between the products individuals buy, which is also known as market basket analysis.

If a buyer purchases ‘A’, would she or he is compelled to buy ‘B’ because of a relationship between the products ‘A’ and ‘B’. This leads to relationship that exists between cheese and cokes etc, while purchasing in a mall. When new products promotions are done in the market, after knowing these relationships it develops new sequence of association. Prior knowledge of these relationships could help in proposing the associated product to the customer. This may likely to increase the chance of the customer buying it; it can also help in combining products for and enhance the sale of products.

This learning of associations between products by a machine is learning associations. Once we found an association by examining a large amount of sales data, Big Data analysts. It can develop a rule to derive a probability test in learning a conditional probability.

6. Classification with respect to ML: It is way of categorising individual for the study. Here individuals are taken from population in distinct groups. By this way variables can be identified independent.

In an banking sector, if a person wants to apply for a loan, before a bank agrees to disburse a loan, Bank will ensure

that customers possesses the ability to reimburse the loan. For this process many factors are considered such as customer’s salary, age, investments and financial background, previous loan status. This information taken from the past data of other loan departments, banks. Hence, It helps to create relationship between customer attributes and risks associated with it.

7.Opinion mining: It is also known as **sentiment analysis or emotion AI**, It refers to the use of text analysis, computational linguistics, natural language processing, and biometrics to systematically identification, extraction, quantification, and study. Sentiment analysis is used widely now a day. It is implemented to voice of the customer, for example reviews of the customer and survey responses, online review of product, social media likes and dislikes.

Sentiment analysis aims to define the attitude and behaviour of a presenter, author, or other subject with respect to some topic or the overall contextual polarity or emotional reaction to a document, interaction, or event. The attitude may be a decision wise or assessment, emotional state, or the intended emotional message. Very Popular example of sentimental analysis is found in Tivago– a website made for comparing the rates of different hotels throughout the country. It takes sentiments in form of smilies of three different colors ranging from red smiley (for poor rating/ unsatisfied customer), orange (for average rating/satisfied customers), and green smiley (for excellent rating/ Happy Customers).

7. Prediction (Forecasting):As the name suggests it gives us the forecasting or prediction. Based on the data collected during the process of ML, the ,model selected for the process the result is predicted. Prediction is likely to be the near to the exact result if the model, process, and hypothesis all are near to correctness.

There are many cases when a business or product need prediction before the actual process starts.

One such case is seen in banks during loan repayment consideration. If banks need to be on safer business, bank will predict whether a person is likely to repay the entire amount or he/she may fall under defaulter category. It can be defined by a set of guidelines set by the predictors. Once the process of prediction is done as per requirement probability can be computed.

8. Extraction: Another popular application of machine learning is extraction. As the name defines itself, it is the process of extracting structured information from unstructured data. There are many examples to name a few are blogs, web pages, articles, business reviews and reports, and e-mails. The RDBMS maintains the output which we can get by the information extraction.

Big data is the today’s buzzword. Data collected from almost all sources contribute to the big terminology i.e.; BIG Data. The term big itself defines something huge, with volume, variety, varacity and velocity. Huge volume of data which is getting generated among it most of the data is unstructured. Now here comes the big challenge with big data to handle unstructured data. Some of the data collected so far can be converted to structured data.

Now a days the way in which data is being collected is also changed, in earlier days data was collected in batches like End-of-Day (EOD), but now business demands of accessing data and analysing is totally changed, they want the data as soon as it is getting generated, i.e. in real time.

9. Predicting Rating:- A recommender system is a subclass of information filtering system that helps in predicting the "rating" or "preference" that a user would give to an item.[12][13]

This has become widespread in present years, and the recommender system is used in a range of areas some of which includes cinemas, songs, news casts, social tags, books, research areas, search requests, and products in general. Many such examples have been seen in restaurants, online shopping sites, fashions, financial services, life insurance, Twitter pages and many more.

10. Business Application: When it comes to business applications of machine learning, marketing is always near the top of the list. Modern digital marketing offers a huge volume of quantifiable data for teams to work with, and marketing can be said to take precedent over other areas like customer service and business intelligence because of its direct tie to driving revenue. Machine learning marketing applications are still relatively novel for most small and medium-sized business, but this may change drastically over the next five years.

Conclusion

In this literature review paper we have seen the machine learning- an incredible innovation in the field AI. Machine learning we can say has improved the lives of human being in term of their successful implementation and accuracy in predicting results. In this paper we have seen some of the successful implemented application

REFERENCES:

- 1) Joseph A. Cruz, David S. Wishart Applications of Machine Learning in Cancer Prediction and Prognosis.
- 2) Guang-BinHuangQin-YuZhuChee-KheongSiew
- a. Extreme learning machine: Theory and applications.
- 3) <https://www.datasciencecentral.com/profiles/blogs/top-9-machine-learning-applications-in-real-world>
- 4) <https://machinelearningmastery.com/a-tour-of-machine-learning-algorithms/>
- 5) <https://towardsdatascience.com/types-of-machine-learning-algorithms-you-should-know-953a08248861>
- 6) https://en.wikipedia.org/wiki/Machine_learning
- 7) <https://www.quora.com/>
- 8) www.expertsystem.com/machine-learning-definition/
- 9) <https://pythontips.com>
- 10) <http://books.tarsoit.com/Machine%20Learning%20with%20R%20-%20Second%20Edition.pdf>
- 11) <https://www.usenix.org/system/files/conference/osdi16/osdi16-abadi.pdf>
- 12) www.mssbti.com
- 13) www.lib.bioinfo.pl
- 14) www.datasciencecentral.com, 2017