A DETAILED STUDY ON MACHINE LEARNING AND IT'S ALGORITHMS

Ratna Dey[1], Aisik Bisoi[2], Sujata Kundu[3], Shyamapriya Chowdhury[3], Soumya Bhattacharyya[3]

[1] Student, B Tech in Information Technology, Narula Institute Of Technology, Kolkata-700109,
E-mail-ratnadey1704@gmail.com
[2] Student, B Tech in Information Technology, Narula Institute Of Technology, Kolkata-700109,
[3] Assistant Professor, Department of Information Technology, Narula Institute Of Technology, Kolkata-700109

Abstract— In today's world our life is totally moved around the data, like IOT, cyber security , mobile information, social, business and data and many more data. We need Artificial Intelligence (AI) to analyze these data and built a powerful automated application and to do these things we need machine learning. Basically, it is a key aspect of AI. ML Algorithms like supervised, reinforcement, unsupervised and semi-supervised learning algorithms are mainly used to analysis the data. In these paper, we have discussed a thorough point of view on these algorithms that can help us in future to apply in suitable instructive assortment to get smart applications for real world.

Keywords— *Machine learning, Artificial intelligence, Supervised Machine Learning, Unsupervised Machine Learning, Reinforcement Super learning, Semi supervised Machine Learning*

I. INTRODUCTION

Machine learning is a part of Artificial Intelligence (AI) which is capable to allow software applications to predict more accurate outcomes by input and analyze large amount of data. Example of ML is image processing, prediction classification, learning association, regression.

In simple word ML gives permission the user to provide a computerized algorithm with vast amount of data to the analysis the data and predict the accurate output if any correction is needed,

Machine learning algorithms are used to correct the codes and provide us a good model as output.

There are basically 4 type's of algorithms:

1. **supervised learning**
2. **unsupervised learning**
3. **reinforcement learning**
4. **semi supervised learning**

When to Use Machine learning???

ML is not the solution for every type's of problem, It is a technique that used only for robust types of problems where human cannot solve it easily by them self like large computation, Machine learning solutions are effectively solve this problem.

II. ARTIFICIAL INTELLIGENCE

Machine learning is a simple process to archive AI. It is a process of human intelligence in machines that is programmed to think and work like human begins minimize their works. The best characteristics of AI is a rational model and with the help of this we can the best change to archive the goals.

AI based invoice management system is a example of AI.

III. SUPERVISED LEARNING

Supervised learning is a machine learning algorithm. We have an input variable (X) and an output variable (Y). We use an algorithm to learn the function of matching input and output.

So, $Y = f(x)$

40

The main goal of is that the mapping function is sufficiently accurate. Given a new set of inputs (X), it can predict the output variable (Y) of the data set.

This is called supervised learning because we know a lot about the input (learning data) and output data for that given dataset and we have to make a predictive model that gives us the appropriate outcomes for a newly fed data.
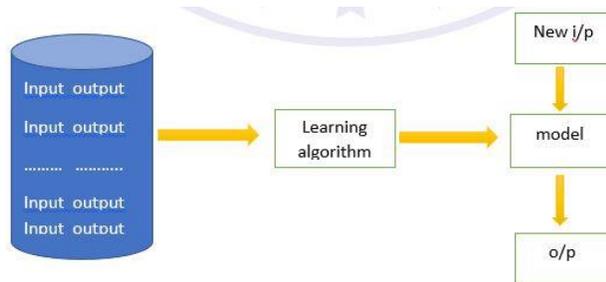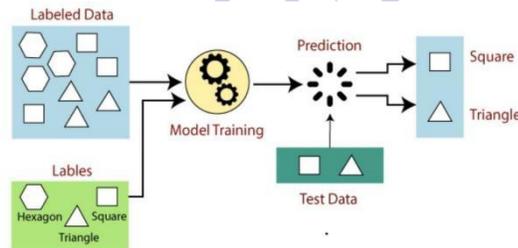

Diagram of supervised algorithm

Example:

A real life example of supervised learning is exit poll. During vote many channel gives exit polls for election.in exit poll they ask peoples for their opinion that which party wins. And after result we compare the prediction that how much it accurate. So here for supervised model the training data is the total predicting of before vote result and the output is actual result after vote .and when we the new input data is the total vote during voting time .if new input and the give output is nearly accurate then it is supervised learning.

Steps for supervised algorithm**:**

1. We have to determine the type of the training data set like square, rectangle, triangle or polygon.

2. Collect the labeled training data.

3. Split it into 3 data sets, such as train data set, test data set and validation data set.

4. Determine the input attributes of the training data set with the largest amount of model data to predict the output.

5. Apply the appropriate algorithm as per the model, for example, support vector machine, choice tree like that.

6. Apply the algorithm to the streaming data set and now and again we need approval sets as the control boundary.

7. At last by providing the test dataset we have to evaluate the accuracy of the predicting model. On the off chance that the model predicts right yield which implies the model is exact.

Basically, there are 2 types of supervised algorithm:

1. <u>Regression:</u>

This type's of algorithm is used when there is connection among input and output variable. It is utilized for expectation the persistent variable like climate estimating, market patterns and so forth. There are various sorts of regression algorithms in supervised learning. They are:

- o   Linear Regression
- o   Regression Tree
- o   Non Linear Regression Tree
- o   Bayesian Linear Regression
- o   Polynomial Regression

2. **<u>Classification:</u>**

When the output is a categorical variable this algorithm is used, so there are 2 cases, for example yes-no, male-female and true-false and so forth.

<u>Supervised learning's Advantages:</u>

1.It is used to sort out different types of real-world-computation problem like fraud detection.

2. We can test and debug the machine.

3.the result is more accurate and reliable than unsupervised learning.

4.We can get the exact idea of classes and objects.

5.The result is highly accurated.

**<u>Supervised learning's Disadvantages :</u>**

1. It is not suitable for to handle the complex tasks.

2. It takes lots of computation time.

3. Always we need to update it.

4. It can be easily over fitted in the model.

IV. UNSUPERVISED LEARNING

Unsupervised learning is another one types of machine learning algorithm. Unsupervised learning is where you just have input information (X) and no comparing output.

The principle is to display the fundamental design or appropriation in the information to get familiar with the information.

This is called unsupervised learning because unlike previous supervised learning there is no real learning. The majority of machines are unsupervised.
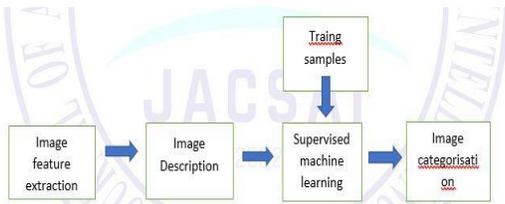
Diagram for unsupervised algorithm

K-means for clustering problem is a prior algorithm.

For example, a dog and a cat, we can identify them by their features (like shape, tail, sound, face difference etc.). We know how a dog or a cat looks like. By their features we make a group or a cluster and put them in the dog or cats category. This is unsupervised learning, where we not taught but we learn from the data.
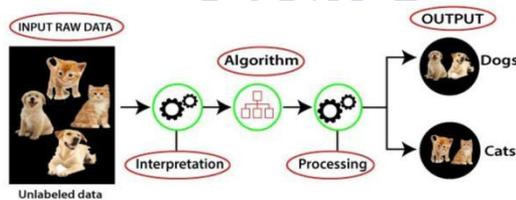
Diagram of unsupervised algorithm example

Steps for unsupervised algorithm :

There are 7 steps for unsupervised algorithm, here are the steps:

1. Data collection
2. Processing of the data
3. Model selection
4. Training
5. Assessment
6. Hyperparameter adjustment
7. Prediction

It is divided into 2 parts, such that:

a. Clustering:

It is a part of unsupervised learning problem which is used to group the unlabeled data set. It is often used as a data analysis technique to discover interesting patterns in data. Eg. Group of customers depends upon by their purchasing behavior.

b. Association:

It is also a part of unsupervised learning problem by which one can discover rules to describe large portion of the data.

Unsupervised learning's Advantages:

1.It is used for more complex tasks as compared to supervised learning because, in unsupervised learning the input data is unlabeled.

2. It is very useful for finding patterns in data which is not possible by using normal methods.

3. As unsupervised learning has unlabeled data set the problems is easy to solve by learning the data classifying it without any labels.

Unsupervised learning's Disadvantages:

1. The result is less accurate because of unlabeled input data.

2. As unsupervised learning does not have output corresponding to the given input, it is more difficult.

3. As the learning phase take lots of time, so it is time consuming.

V. REINFORCEMENT LEARNING

In simple words, It is another type of machine learning in which a programmer interacts with the environment and act within that. For each and every good action it gives +VE feedback and for every bad action it gives negative feedback. As there is no label data the programmer have to learn by its experience. For specific types of problem where decision making is sequential and which have long-term goal, the Reinforcement learning is used like robotics, game playing.
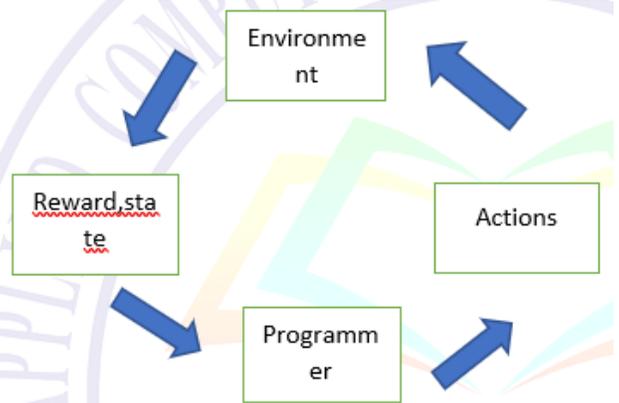
Diagram of Reinforcement learning

Chess game is the good example of Reinforcement Learning.

Some terms in Reinforcement Learning:

  i.   **Agent ()**
 ii.   **Environment ()**
iii.   **Action ()**
 iv.   **State ()**
  v.   **State ()**
 vi.   **Reward ()**
vii.   **Policy ()**
viii.   **Value ()**
 ix.   **Q-value ()**

There are two type's of model in this part:

a) **positive reinforcement learning:**

42

It is occurred due to increasing the particular behavior and strength and the frequency of the behavior.

Advantages**:**

It maximizes performance.

Disadvantages**:**

Too much of positivity may lead to overloaded state that diminishes the output.

b) negative reinforcement learning:

It is opposite to the positive learning. Depending on the strength and behavior, It may be more efficient than positive reinforcement learning which may reduce consequences.

Advantages:

It increases behavior.

Disadvantages:

It affords sufficient to fulfill up the minimal behavior.

Approaches for Reinforcement learning:

1.Valued Based

2.policy based

3.Model-based

VI. SEMI SUPERVISED

It is the blend of Supervised learning and Unsupervised learning. It utilizes limited quantity of labeled information and enormous measure of unlabeled information to prepare a ML model.

Example:

The text document classifier is case of semi supervised learning. as a person don not have the efficient time to read the entire text document, it would be practically difficult to track down a lot of labeled text documents.

STEPS TO WORK:

1. First, use small amount of supervised learning on the dataset to get a good result.

2. Then use the unsupervised learning on dataset for predicting outputs. IT is the pseudo label.

3. Then link the supervised learning and unsupervised learning input dataset.

4. To improve the accuracy and decrease the error, train the data model with supervised learning method as the first step.

VII. CHOOSE THE RIGHT MACHINE LEARNING MODEL

This is the following steps:

Step 1:

Consider a problem with prospective data inputs which can be considered for the solution. This step is done by data scientist who has deep understanding about this.

Step 2:

Collect the data from dataset format it and labeled it.

Step 3:

First choose a algorithm and test the dataset to see how well they perform.

Step 4:

Continue to find the outputs with until they give acceptable accuracy label.

Applications of Machine learning algorithms:

Here's some applications that use machine learning:

1.Automatic Recognition of Hand written postal codes.

2.Computer Aided Diagnosis.

3.Computer version

4.Automatic cars

5.Face recognition

6.speech recognition

7.Data mining

Machine Learning's Future:

Now a day machine learning is on high demand. All the big projects and technical thing and also robotics we use machine learning and all software companies and all top MNC are working on machine learning to improve the system and make easy our life by built the smart and intelligence applications.

ADVANTAGES OF MACHINE LEARNING:

1.ML algorithms are capable to learn from the provided data. Though data increases day by day, the model efficiency and accuracy is increasing day by day.

2. The main utility of ML is It has ability to do various decision making task. ML is changing the world by its automation technique for maximum things that we can think of.

3. One of the most useful advantage is pattern recognition like pattern checker, searching engine of Google.

DISADVANTAGES OF MACHINE LEARNING:

1. Data acquisition is one of the main disadvantages of machine learning and data science. Data collection form all source might be contain huge incorrect data.

2. As ML models are more capable to process huge amount of data, the time to learn and process data is increases. So it is time consuming.

3. To solve machine learning problems we can implement various algorithms manually to choose which algorithm is best for which model.

Future scope Of Machine Learning:

Now a day's machine learning is on high demand. All the big projects and technical thing and also robotics we use machine learning and all software companies and all top MNC are working on machine learning to improve the system and make easy our life by built the smart and intelligence applications.

Conclusion:

We have discussed different machine learning algorithms, its applications and their future scope which helps us to deal with this real word problems.

REFERENCES

[1] Durant, W. Is man a machine? in The Pleasures of Philosophy: A Survey of Human Life and Destiny. New York: Simon and Schuster, 1953, p. 60.

[2] Campbell, M., Hoane Jr. A. J., and Hsu, F.-H. Deep blue, Artificial Intelligence 134(1–2): 57-83, 2002.

[3] journal paper : Machine Learning Algorithms Applications

[4] Talwar, A. and Kumar, Y., 2013. Machine Learning: An artificial intelligence methodology. International Journal of Engineering and Computer Science, 2, pp.3400-3404.

[5] Muhammad, I. and Yan, Z., 2015. Supervised Machine Learning Approaches: A Survey. ICTACT Journal on Soft Computing, 5(3).

Authors Profile

Author Ratna Dey, pursing Btech from narula Institute of technology, Dept of IT ,4th year.

Author Asik Bisoi, pursing Btech from narula Institute of technology, Dept of IT ,4th year.

Author Sujata Kundu received her B.Tech in Information Technology from BIT,WBUT and M.Tech in Computer Science & Application from Calcutta University. Currently she is working as Assistant Professor in the Dept. of Information Technology at Narula Institute of Technology, Agarpara. Her research interest includes IOT, data science, machine learning and deep learning.

Author Shyamapriya Chowdhury received her B.Tech in Information Technology from BPPIMT and M.Tech from BESU, Shibpur. Currently she is working as Assistant Professor in the Dept. of Information Technology at Narula Institute of Technology, Agarpara. Her research interest includes machine learning IOT.

Author Soumya Bhattacharyya received his B.Tech in Information Technology from AIEM,WBUT and M.Tech in Software Engineering from JISCE, Kalyani, Makaut. Currently he is working as an Assistant Professor in the Dept. of Information Technology at Narula Institute of Technology, Agarpara. His research interest includes IOT, data science, machine learning and deep learning. He is also a member of CSI Kolkata Chapter.

44