Machine Learning Algorithms: Analysis and Application

Arpit Kakkar UID: 20MSM3060

Department of Mathematics, Chandigarh University

Abstract: In this paper we will discuss various machine learning algorithms with their applications like image processing, data mining, predictive analysis, natural language processing, etc. These algorithms use programmed instructions that receive and analyse input data to previse the outputs. Every time the new data is fed as an input to these algorithms and they optimize the output in terms of performance and efficiency. The advantage of using machine learning algorithms, it works automatically after learning from the steps to deal with data and gives the desired output.

Introduction: The term Machine Learning often confused with Artificial Intelligence, but it is the sub-domain of AI. Machine learning is a predictive modelling that enables computer to handle data and perform analysis. With huge amount of data stored its impossible to extract data from the dataset and interpret the pattern, in that case, we use machine learning algorithms to find statistical regularities and other patterns from the data. Nowadays machine learning is in very high demand and many industries from medical to the military apply these algorithms to extract the relevant insights from the datasets.

The method of data analysis that automates the process of analytical model building by learning from the data is machine learning. It is a sub-domain of artificial intelligence used to get insights from the past or live data and make decisions via predictive modelling techniques. Computational learning theory is a branch of statistics that deals in the performance and computational analysis of machine learning algorithms.

The primary goal is to train or allow the computer to learn automatically with less human assistance and take the decisions accordingly. A huge amount of data can be analyzed using machine learning algorithms. These algorithms are faster and accurate, requires additional time and resource to get trained properly. We can combine these algorithms with AI to make it more effective in processing a very large amount of data.

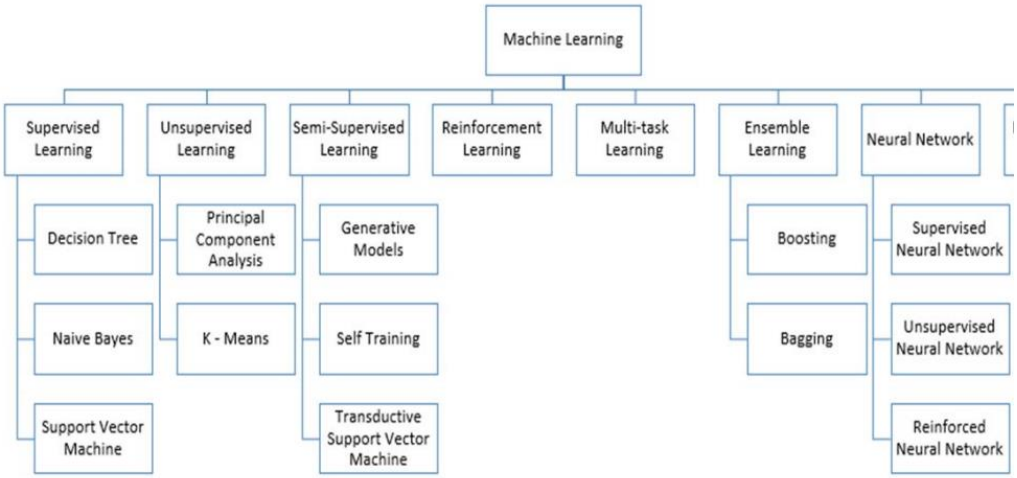

Fig. 1. Types of Machine Learning Algorithms

Machine learning algorithms are often divided into two categories i.e. supervised and unsupervised learning, but there are other algorithms too that can play a major role in machine learning. We will discuss all the techniques in the coming sections.

1. Supervised Learning

Supervised learning algorithms are those that need external assistance. As the name suggests these algorithms are designed to learn by example, it is like a supervisor or teacher to train the whole process. The set of inputs with pre-defined outputs is provided as training data and the algorithm will search for the patterns in the data. After training, the algorithm is implemented to the new set of data with new inputs that will be classified. The objective of the supervised learning algorithm is to predict the outputs for new data.

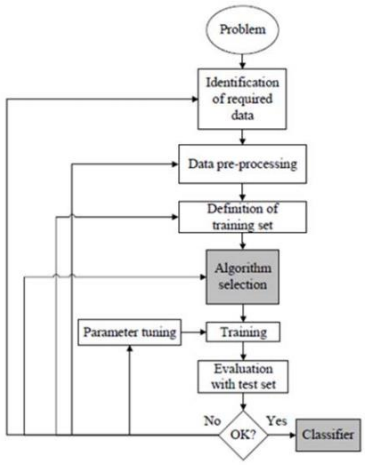

Fig. 2. Workflow of Supervised machine learning algorithm

- **Decision Tree:** A decision tree is also referred to as a classification tree or a reduction tree. As the name suggests, it uses a tree-like model to specify sequences of decisions and consequences. The decision tree uses predictive models to predict the behavior that has not been tested. I.e. if an organization wants to switch from an analog controller to a digital controller, this model can be used to test the performance change. In a decision tree predictive model is used to map the observations about an item to a conclusion about its output value. The response predictions can be achieved by making a decision tree with testing points and branches.

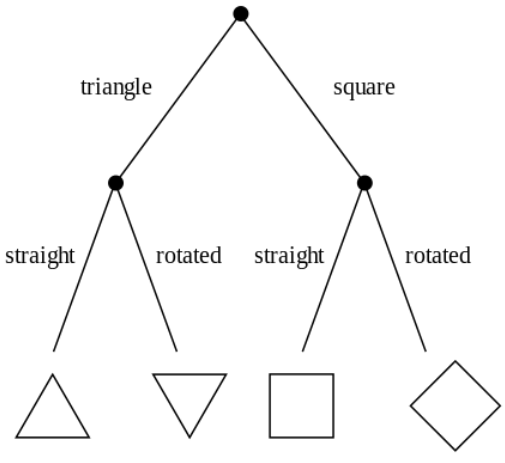

Fig. 3. Simple Decision Tree

- **Naive Bayes:** This algorithm is based on Bayes Theorem which follows probabilistic technique and mainly used for clustering and classification purpose. But it works well for natural language processing problems. Naïve Bayes creates trees based on the conditional probability of happening and these trees are known as Bayesian Network.

- **Support Vector Machine:** It is another widely used machine learning technique and mainly used for both regression and classification tasks. It works on the principle of margin calculation. The main objective of this algorithm is to find an N-dimensional hyperplane. In this technique, we choose hyperplane to separate the two classes of data points and find a plane which has the maximum margin. (i.e. margin is the maximum distance between the data points of both the classes) and these decision boundaries of hyperplane is to classify the data points. Data points which falls on either side of the boundaries are considered as the different classes.

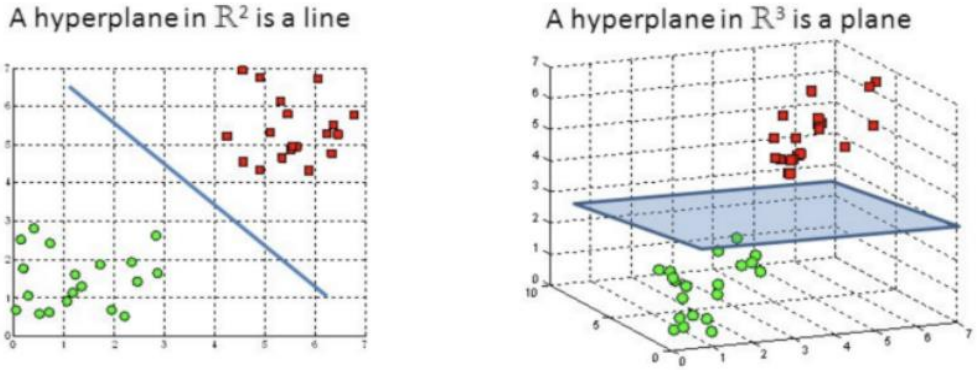

Fig. 4. Hyperplane in 2D and 3D Space

2. Unsupervised Learning

Like supervised learning algorithms, unsupervised learning algorithms do not need a supervisor or teacher to train the whole process. It uses unlabelled datasets and works on its own and discovers hidden insights or patterns that were previously not discovered. The more complex task can be performed using unsupervised learning as compared to supervised learning. Unsupervised learning problems are further categorised into clustering and feature reduction problems. Clustering is well known unsupervised learning problem.

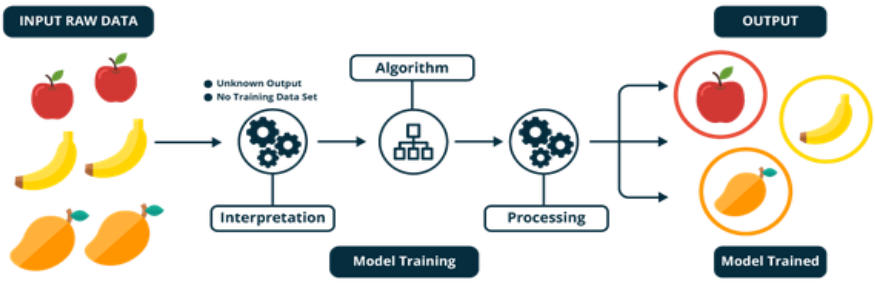

Fig. 5. Example of Unsupervised Learning

K-Means Clustering and Principal Component Analysis are two main algorithms for clustering and feature reduction comes under unsupervised learning.

- **K-Means Clustering:** A collection of similar data objects is known as clusters and the process of organizing or collecting those data objects into groups is clustering. The objects having same properties grouped together and put in the same cluster. This algorithm creates K distinct clusters; the center of the cluster contains the mean of the values in a particular cluster. The below figure shows the un-clustered data and clustered data.

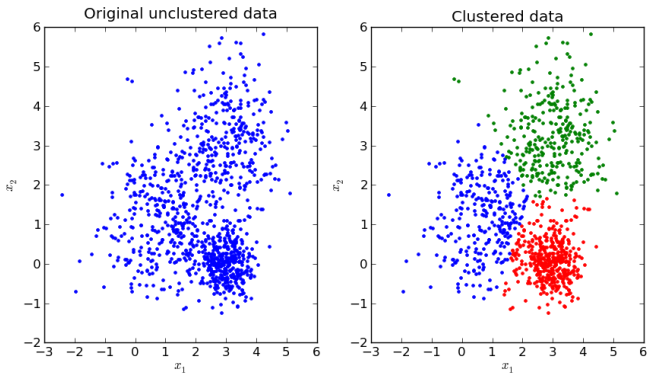

Fig. 6. K-Means Clustering

- **Principal Component Analysis:** Some datasets contain a very high dimensional nature which makes direct visualization impossible. The technique used to reduce the data dimension is known as Principal Component Analysis or PCA. It is a well-known mathematical technique in which dimension reduction can be achieved by creating new artificial variables called principal components. To understand the PCA, we take an example of 2D data. Before PCA it takes two axes and after applying PCA the data is plotted in 1D.

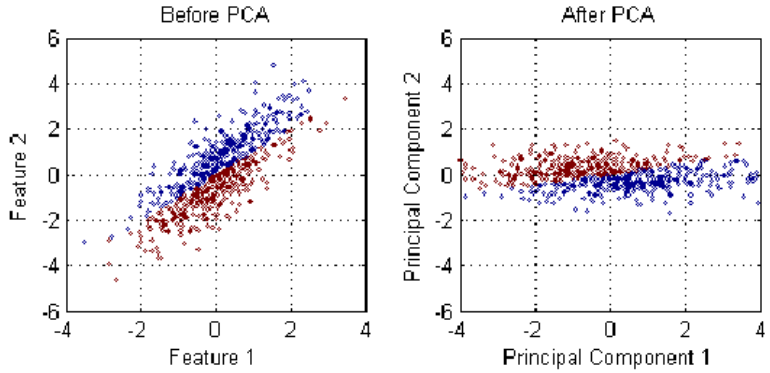

Fig. 7. Visualization of data before and after PCA

3. Semi-Supervised Learning

The semi-supervised learning algorithms are the mixture of supervised and unsupervised learning techniques. This algorithm can be trained using the dataset that contains both labeled and unlabelled data, but mostly unlabelled data is preferred. We use this technique when we don't have enough labeled data to produce an accurate result and neither we have resources to get labeled data.

- **Generative Models:** Generative models is one of the oldest technique used in semi-supervised learning. It processes a large amount of training data and data reduction is

performed digitally and generates new data instances. These models are implemented on neural networks and generate new reduced data instances.

- **Self-Training:** In self-training models, the classifier is trained using a small amount of labeled data and after training it is provided with the unlabelled data to make the predictions. The classifier starts learning itself when this process is repeated many times, hence known as the self-training technique.
- **Transductive Support Vector Machine:** It is an extension of SVM and widely used for treating partially labeled data. But labeled and unlabelled data both can be considered. It can be used to label the unlabelled data and the margin between them is maximum.

4. Reinforcement Learning

Reinforcement learning is a technique in which an agent is bound to learn from its own actions and experiences. The agent has no prior knowledge of which actions to take, it uses the trial and error method to get the best outcome. Unlike supervised learning, no feedback is provided at the end for the correct set of actions in reinforcement learning. It uses reward and punishment as a signal for positive and negative behavior. It requires a huge amount of data to perform well, so this learning technique is mostly implemented where simulated data is available like games and robotics.

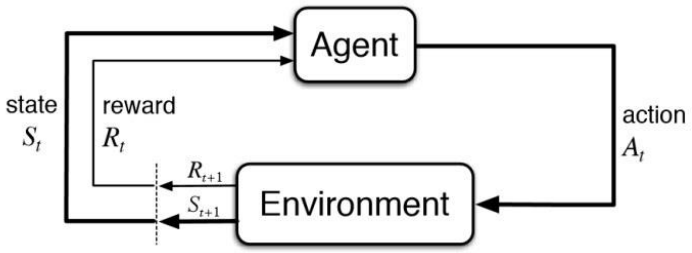

Fig. 8. The Reinforcement Learning Model

5. Multi-task Learning

Multi-task learning is associated with deep learning and concepts of neural networks are widely used in it. The main goal of this learning is to help other learners to perform better. This algorithm remembers the steps and procedures it uses to solve the particular problem and how it reached the conclusion and later it uses the same steps and procedure to find the solution of other similar tasks. If learners share their experiences with each other they can perform the task much faster. The multi-task algorithm can be used in a variety of applications like speech recognition, natural language processing, drug discovery, etc.

6. Ensemble Learning

In ensemble learning algorithms, multiple individual learners are trained and these individual learners are called ensemble members. The outputs of these ensemble members are combined into a single prediction and this final output has better performance as compared to individual learners. These ensemble members may be naïve Bayes, decision trees, neural networks, etc. There are two types of ensemble learning techniques given below.

- **Boosting:** Boosting is a sequential ensemble technique that transforms weak learners into strong learners. It decreases the bias and gives very strong predictive models. The primary goal is to achieve high accuracy by converting weak learners into better-performing models. The best example is Ada Boost or Adaptive Boosting, that helps combining the multiple weak classifiers into a single strong classifier.

Input: Data set $\mathcal{D} = \{(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2), \cdots, (\boldsymbol{x}_m, y_m)\}$;
Base learning algorithm $\mathcal{L}$;
Number of learning rounds T.

Process:
$D_1(i) = 1/m$.
for $t = 1, \cdots, T$:
$h_t = \mathcal{L}(\mathcal{D}, D_t)$;
$\epsilon_t = \Pr_{i \sim D_i}[h_t(\boldsymbol{x}_i) \neq y_i)]$;
$\alpha_t = \frac{1}{2} \ln \frac{1-\epsilon_t}{\epsilon_t}$;

$$D_{t+1}(i) = \frac{D_t(i)}{Z_t} \times \begin{cases} \exp(-\alpha_t) & \text{if } h_t(\boldsymbol{x}_i) = y_i \\ \exp(\alpha_t) & \text{if } h_t(\boldsymbol{x}_i) \neq y_i \end{cases}$$
$$= \frac{D_t(i) \exp(-\alpha_t y_i h_t(\boldsymbol{x}_i))}{Z_t}$$

end.

Output: $H(\boldsymbol{x}) = \text{sign}\,(f\,(\boldsymbol{x})) = \text{sign}\ \sum_{t=1}^{T} \alpha_t h_t(\boldsymbol{x})$

Fig. 9. Pseudo code for AdaBoost

- **Bagging:** It stands for Bootstrap Aggregating. It is a parallel ensemble technique and the primary goal is to reduce the variance of the prediction model by forming additional data at the time of training.

Input: Data set $\mathcal{D} = \{(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2), \cdots, (\boldsymbol{x}_m, y_m)\}$;
Base learning algorithm $\mathcal{L}$;
Number of learning rounds T.

Process:
for $t = 1, \cdots, T$:
$\mathcal{D}_t = Bootstrap(\mathcal{D})$;
$h_t = \mathcal{L}(\mathcal{D}_t)$
end.

Output: $H(\boldsymbol{x}) = \text{argmax}_{y \in \mathcal{Y}} \sum_{t=1}^{T} 1(y = h_t\,(\boldsymbol{x}))$

Fig. 10. Pseudo code for Bagging

7. Neural Network Learning

It is also known as artificial neural network or ANN and the concept of neural network is derived from biology where neurons are a cell-like structure in a brain. To fully understand the working of neural networks we have to know how neurons work in our brain. A neuron consists of mainly 4 parts viz. dendrites, nucleus, axon, and soma.

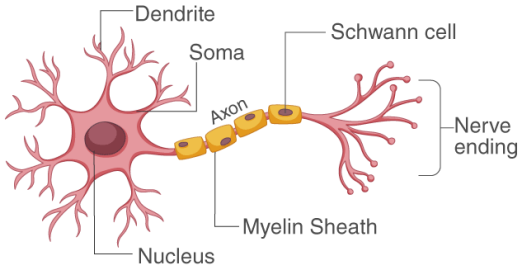

Fig. 11. The Structure of Neuron

An electrical signal is received by the dendrites and processed by the soma. Axon carries the output of the processed signal to the dendrites terminals and then it is transferred to the next neuron. The heart of the neuron is a nucleus. The electrical impulse travels around the brain through neurons and this interconnection of neurons is known as a neural network.

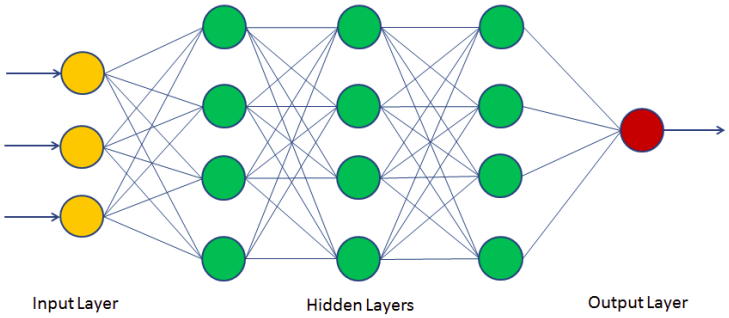

Fig. 12. Structure of Neural Network

An artificial neural network works like our brain. There are mainly three layers viz. input layer, hidden layer, and output layer. The input layer takes input and hidden layer process that input and output layer then sends the processed output. The neural network can be divided into three basic categories: supervised, unsupervised, and reinforced.

- **Supervised Neural Network:** The concept of supervised neural networks is the same as supervised learning algorithms; the output is already known for the set of inputs. The predicted outputs are then compared with the know outputs and the parameters have been changed based on the variations in the outputs. With changed parameters, input is again provided to the neural network to get the desired outputs.

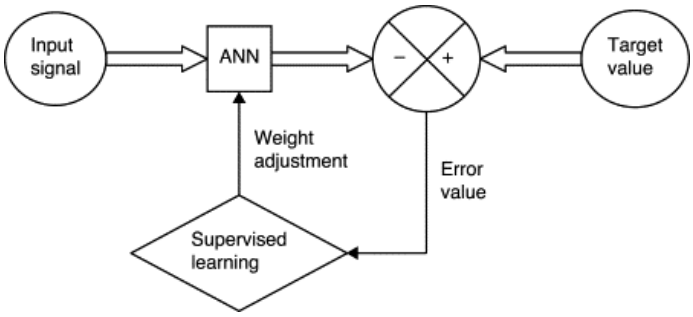

Fig. 13. Supervised Neural Network

- **Unsupervised Neural Network:** It is the same as the unsupervised learning algorithm; no idea about the output at the beginning. The network categorizes the data as per the similarities between them. The grouping can be done by the neural network after checking the correlation among the inputs.
- **Reinforced Neural Network:** It is the same as the reinforcement learning algorithm; agent interacts with the environment as we human do and some reward has been provided on the basis of decision taken by the network. If the decision taken is correct then the connection between the corresponding points is strengthened and weakened if the decision is wrong. In this way, the reinforced neural network works and it has no prior information about the outputs.

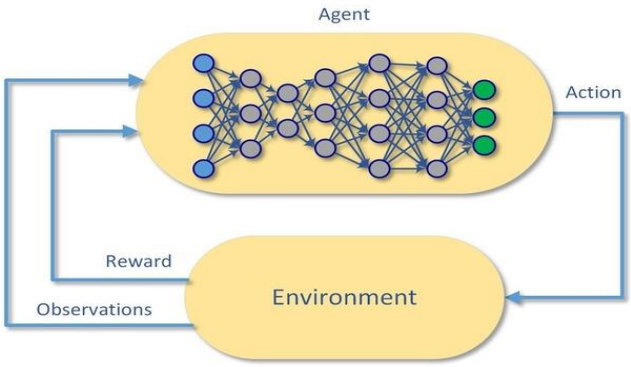

Fig. 14. Reinforced Neural Network

Applications of Machine Learning Algorithms

1. **Image Recognition**
   It is the most common application of machine learning and used to identify persons, objects, places, etc. Traditional algorithms line K-Nearest Neighbour (KNN) and Support Vector Machine (SVM) is used for image classification and recognition. The most popular use case of face detection is automatic friend tagging suggestions given by social media platforms like Facebook. It is based on the Python library and the name of the framework is deepface and used by Facebook under the project named "Facebook DeepFace". This library is based on Keras and Tensor Flow, a python deep learning API used in machine learning.

2. **Speech Recognition**
   While using smartphones we get the option to search by voice option, it comes under speech recognition and very popular machine learning application. It uses a supervised learning algorithm and deep neural network techniques. It can enable the program to process human speech into written text and search accordingly. There is a python library named SpeechRecognition that works with the support of APIs provided by Google, IBM, Microsoft, etc.

3. **Recommender Systems**
   Machine learning is widely adopted by various e-commerce, entertainment, and OTT companies like Amazon, Flipkart, Netflix, etc. to provide product recommendations to the users. It is based on deep learning and neural network and analyses the user search and purchase patterns. It uses a collaborative filtering technique to filter out the data.

4. **Medical Diagnosis**
   Machine learning is widely used in the field of medical science for disease diagnoses. One of the best examples is a prediction of potential heart failure, an algorithm is designed to scan and identify the patterns in a patient's cardiovascular history and making the analysis using the medical reports, and no need for a physician to dig through multiple health records.

Conclusion

This paper covers popular machine learning algorithms used by many data scientists. Nowadays every person taking help from these machine learning algorithms willingly or unwillingly. From getting recommendation systems in online shopping or OTT platforms are using these algorithms to get the things automated. The core idea of this paper is to give a basic idea about the commonly used machine learning algorithms.

References

1. https://machinelearningmastery.com/a-tour-of-machine-learning-algorithms/
2. https://www.expert.ai/blog/machine-learning-definition/
3. https://towardsdatascience.com/a-brief-introduction-to-supervised-learning-54a3e3932590
4. https://www.kdnuggets.com/2020/06/naive-bayes-algorithm-everything.html
5. https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47
6. https://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/
7. https://towardsdatascience.com/understanding-k-means-clustering-in-machine-learning-6a6e67336aa1
8. https://towardsdatascience.com/unsupervised-learning-and-data-clustering-eeecb78b422a
9. https://towardsdatascience.com/a-one-stop-shop-for-principal-component-analysis-5582fb7e0a9c
10. https://www.kdnuggets.com/2012/11/qlucore-understanding-principal-component-analysis.html#:~:text=Principal%20Component%20Analysis%20(PCA)%20is,artificial%20variables%20called%20principal%20components.
11. https://towardsdatascience.com/supervised-learning-but-a-lot-better-semi-supervised-learning-a42dff534781
12. https://scikit-learn.org/stable/modules/neural_networks_supervised.html
13. https://medium.com/intuitive-deep-learning/autoencoders-neural-networks-for-unsupervised-learning-83af5f092f0b