Data Analytics and Artificial Intelligence
Vol: 2(1), 2022
REST Publisher; ISBN: 978-81-948459-4-2
Website: http://restpublisher.com/book-series/data-analytics-and-artificial-intelligence/

A Review on Machine Learning

Josephine reena mary .S, Rev. Sr. ArockiaValan Rani
ST. Joseph's College of Arts and Science for Women, Hosur ,Tamil nadu ,India
*Corresponding author Email: josephinemary775@gmail.com

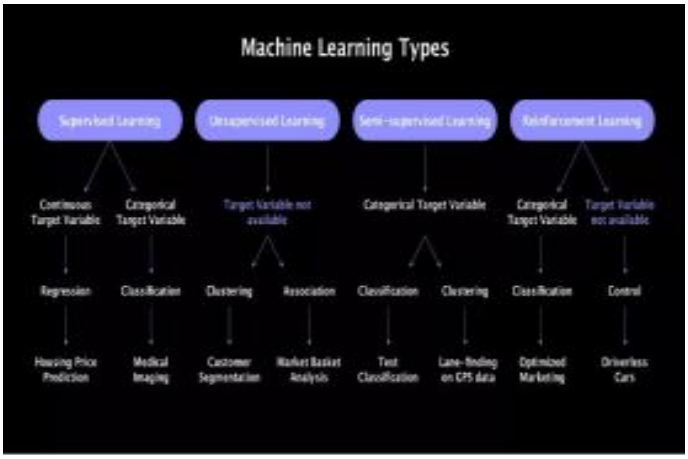
Abstract: Machine learning (ML) is the scientific study of algorithms and statistical models that computer systems use to perform a specific task without being explicitly programmed. Learning algorithms in many applications thats we make use of daily. Every time a web search engine like Google is used to search the internet, one of thereasons that work so well is because a learning algorithm that has learned how to rank web pages. These algorithms are used for various purposes like data mining, image processing, predictive analytics, etc. to name a few. The main advantage of using machine learning is that, once an algorithm learns what to do with data, it can do its work automatically. In this paper, a brief review and future prospect of the vast applications of machine learning algorithms has been made.

Keywords: Algorithm, Machine Learning, Pseudo Code, Supervised learning, Unsupervised learning, Reinforcement learning

1. Introduction

A good start point for this paper will be to begin with the fundamental concept of machine learning. In machine learning a computer program is assigned to perform some tasks and it is said that the machine has learnt from its experience if its measurable performance in these tasks improves as it gains more and more experience in executing these tasks. Sothe machine takes decisions and does predictions forecasting based on data. Take the example of computer program that learns to detect predict cancer from the medical investigation reports of a patient. It will improve in performance as it gathers more experience by analyzing medical investigation reports of wider population of patients. Its performance will be measured by the count of correct predictions and detections of cancer cases as validated by an Since their evolution, humans have been using many types of tools to accomplish various tasks in a simpler way. The creativity of the human brain led to the invention of different machines. These machines made the human life easy by enabling people to meet various life needs, including travelling, industries, and computing. And machine learning is the one among

2. Supervised Machine Learning

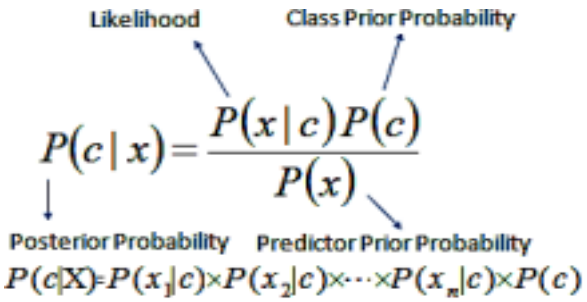


- It is the basic type of Machine Learning Algorithms where the programmer has greater control over the process.
- The engineer can decide which data he or she feeds into the system and which type of output is expected from the system.
- The machine must process the given data and provide solutions in the desired manner.
- The supervised machine learning algorithms are those algorithms which needs external assistance.
- The input dataset is divided into train and test dataset. The train dataset has output variable which needs to be predicted or classified.

3. Decision Tree

Decision tree is a graph to represent choices and their results in form of a tree. The nodes in the graph represent an event or choice and the edges of the graph represent the decision rules or conditions. Each tree consists of nodes and branches. Each node represents attributes in a group that is to be classified and each branch represents a value that the node can take**.**

The naive Bayes algorithm is based the Bayes' theorem with the assumption of independence between each pair of features. It works well and can be used for both binary and multi-class categories in many real-world situations, such as document or text classifcation, spam fltering, etc. The key beneft is that, compared to more sophisticated .approaches, it needs a small amount of training data to estimate the necessary parameters and quickly.

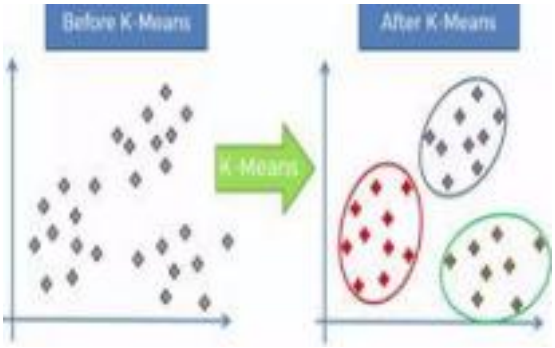


instance x, instance set X concept c ⊆ X, or c : X → { 0, 1 } example (labeled instance): hx, c(x)i; positive examples, neg. examples hypotheses h : X → { 0, 1 } hypotheses representation language hypotheses set H hypotheses consistent with the concept c: h(x) = c(x), ∀ example hx, c(x)i version space. learning = train + test supervised learning (classification), unsupervised learning (clustering) . errorh = | { x ∈ X, h(x) 6= c(x)} | training error, test error accuracy, precision, recall validation set, development set n-fold cross-validation, leave-one-out cross validation over fitting

4. Unsupervised Learning

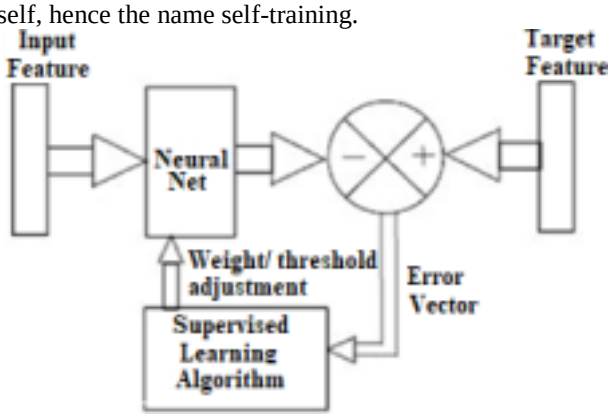
These are called unsupervised learning because unlike supervised learning above there is no correct answers and there is no teacher. Algorithms are left to their own devises to discover and present the interesting structure in the data. The unsupervised learning algorithms learn few features from the data. When new data is introduced, it uses the previously learned features to recognize the class of the data. It is mainly used for clustering and feature reduction. A. K-Means Clustering. K-means is one of the simplest unsupervised learning algorithms that solve the well known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters. The main idea is to define k centers, one for each cluster. These centers should be placed in a cunning way because of different location

causes different result. So, the better choice is to place them is much as possible far away from each other.B. Unsupervised Neural Network the neural network has no prior clue about the output the input. The main job of the network is to categorize the data According to some similarities. The neural network checks the correlation between various inputs and groups them.



5. Semi Supervise Learning

 Semi-supervised machine learning is a combination of supervised and Unsupervised machine learning methods.
 It can be fruit-full in those areas of machine learning and data mining where the unlabeled data is already present and getting the labeled data is a tedious process.
 Common supervised machine learning methods, you train a machine learning algorithm on a "labeled" dataset in which each record includes the outcome information.A. Transductive SVM
 Transductive support vector machines (TSVM) has been widely used as a means of treating partially labeled data in semisupervised learning.
 Around it, there has been mystery because of lack of understanding its foundation in generalization.
 It is used to label the unlabeled data in such a way that the margin is maximum between the labeled and unlabeled data.
 Finding an exact solution by TSVM is a NPhard problem.B. Generative Models
 A Generative model is the one that can generate data.
 It models both the features and the class (i.e. the complete data).
 If we model P(x,y): I can use this probability distribution to generate Data points - and hence all algorithms modeling P(x,y) are generative.
 One labeled example per component is enough to confirm the mixture distribution.C. Self-Training
 In self-training, a classifier is trained with a portion of labeled data.
 The classifier is then fed with unlabeled data.
 The unlabeled points and the predicted labels are added together in the training set.
 This procedure is then repeated further.
 since the classifier is learning itself, hence the name self-training.



6. Conclusion

Machine Learning can be a Supervised or Unsupervised. If you have lesser amount of data and clearly labelled data for training, opt for Supervised Learning. Unsupervised Learning would generally give better performance and results for large data sets.If you have a huge data set easily available, go for deep learning techniques. You also have learned Reinforcement Learning and Deep Reinforcement Learning. You now know what Neural Networks are, their applications and limitations. This paper surveys various machine learning algorithms. Today each and every person is using machine learning knowingly or unknowingly

References

[1]. W. Richert, L. P. Coelho, "Building Machine Learning Systems with Python", Packt Publishing Ltd., ISBN
[2]. 978-1-78216-140-0
[3]. J. M. Keller, M. R. Gray, J. A. Givens Jr., "A Fuzzy K Nearest Neighbor Algorithm", IEEE Transactions on
[4]. Systems, Man and Cybernetics, Vol. SMC 15, No. 4,August 1985
[5]. https://www.geeksforgeeks.org/machine learning/
[6]. S. Marsland, Machine learning: an algorithmic perspective. CRC press, 2015.
[7]. M. Bkassiny, Y. Li, and S. K. Jayaweera, "A survey on machine learning techniques in cognitive radios," IEEE

[8]. Communications Surveys & Tutorials, vol. 15, no. 3,pp. 1136–1159, Oct. 2012.
[9]. https://en.wikipedia.org/wiki/Instance based_learning
[10]. R. S. Sutton, "Introduction: The Challenge of Reinforcement Learning", Machine Learning, 8, Page
[11]. 225-227, Kluwer Academic Publishers, Boston, 1992
[12]. P. Harrington, "Machine Learning in action", ManningPublications Co., Shelter Island, New York, 2012