International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 03 | Mar 2021 www.irjet.net p-ISSN: 2395-0072

Study of Machine Learning Types and Algorithms

Saranya T[1], Revathi S[2]

[1]PG Student, Department of Computer Science, Dr.N.G.P Arts and Science College, Tamil Nadu, India
[2]Assistant Professor, Department of Computer Science, Dr.N.G.P Arts and Science College, Tamil Nadu, India

---

Abstract- *Machine learning is cares with computer programs that automatically improves their performance through experience. It enables IT systems to acknowledge patterns on the thought of existing algorithms, data sets and to develop adequate solution concepts. In this paper, various types of machine learning algorithms are discussed and these algorithms are used for various purposes like data processing , image processing, predictive analytics, etc. to name a few. The main advantage of using machine learning is that, once an algorithm learns what to try to do with data, it can do its work automatically.*

Key Words: Machine Learning, Supervised learning, Unsupervised learning

1. INTRODUCTION

Machine Learning is a one of the sub-area of AI, whereby it refers to the power of IT systems to independently find solutions to problems by recognizing patterns in databases.[1]
Machine learning is used for teach machines how to handle the data more efficiently. Sometimes after viewing the data, we cannot interpret the pattern or extract info from the data. In that case, we apply machine learning[2].
Machine learning models allows computer to learn as the human learn and perform accurately on new and unseen problems after learning from past data of similar problem. Training the system using an efficient algorithm to solve optimization problem with massive amount of data in the form of past experience trained model for unseen problem and improve the performance over the period of time[5] .Focuses on the design and development of algorithm base on experimental data.

2. TYPES OF LEARNING

2.1 SUPERVISED LEARNING

It allows you to gather data or produce a knowledge output from the previous experience. Helps you to optimize performance criteria using experience. Supervised machine learning helps you to unravel various sorts of real-world computation problems used in predication, classification and regression based application

In supervised learning, both the inputs and the outputs can be observed. Here, a supervisor (or teacher) provides sets of (X, Y) pairs, where X denotes input and Y denotes the corresponding output. We are given a training set of examples of two categories C1 and C2. Construct a model that will accurately classify future examples in appropriate categories. This means that in supervised learning, learn good estimation of a function f that transforms given input X into output Y , which has corresponding labelled instances, i.e., using a training set {(X 1 , Y 1), (X 2 , Y 2), ..., (X n , Y n)}, where X = { X i } is input set of examples and Y = { Y i } is the corresponding output along with its category.
There are two major types of problems with supervised learning: classification involving prediction of a class label and regression involving prediction of a numerical value.

Regression: A problem of supervised learning involving predicting a numerical mark.

Classification: Supervised problem of learning that requires the prediction of a class label. They are some classification algorithms are discussed in below.

2.1.1 LOGISTIC REGRESSION

Logistic regression falls under the classification of supervised learning .It is used for the probability estimation of a target variable. The existence of the target or variable is dichotomous, indicating that only two possible groups will be possible. In simple words, with data coded as either 1 (stands for success/yes) or 0 (stands for failure/no), the dependent variable is binary in nature.

Mathematically, as a function of X, a logistic regression model predicts P(Y=1). It is one of the only ML algorithms to be used for multiple classification problems, including such prediction of prediction, Heart disease, prediction of diabetes, cancer detection, etc.

2.1.2 K-NEAREST NEIGHBOR(KNN)

People typically point to solutions to similar problems that have already been solved while attempting to solve new problems. The nearest K neighbor is the classification algorithm in which the decision to assign the new instance to the class or group is made in this process by comparing many K of the most related samples or neighbors in the training set[6].

The number of samples is counted in each class and the new sample is assigned to the groups to which the greater number of neighbors belong. K is a parameter for which the best value must be selected by mutual validation. To find the nearest neighbor, the nearest neighbor needs to define a distance function. The normalization of the mean and division of the standard deviation is the regular form for numerical input. For independent input, Euclidean distance is used[3]. Otherwise, Mehnalubis distance is used. For binary features, Jaccard distance can be used.

As a model, the power of K-'s closest neighbor does not need simple preparation. More data causes higher learning automatically (and old data can be deleted). Although the data must be ordered by the tree and also to find the smallest neighbor with time complexity O (LOG N) that is more than O (N), the drawback of k-NN, on the other hand, is that high dimensions can not be well tolerated.

2.1.3 DECISION TREE

Decision Trees are fall into the Supervised Machine Learning (that is you explain what the input is and what the corresponding output is within the training data) where the info is continuously split consistent with a particular parameter. The tree are often explained by two entities, namely decision nodes and leaves. The leaves are the choices or the ultimate outcomes. And therefore the decision nodes are where the info is split[4].

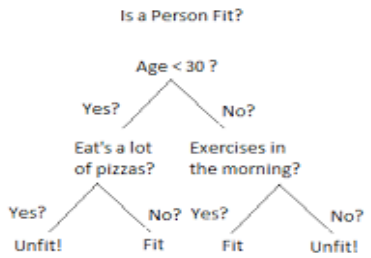

Fig.1 Example of Decision tree

Decision Trees modified An example of a choice tree are often explained using above binary tree. Let's say you would like to predict whether an individual is fit given their information like age, eating habit, and physical activity, etc. the choice nodes here are questions like 'What's the age?', 'Does he exercise?', 'Does he eat tons of pizzas'? And therefore the leaves, which are outcomes like either 'fit', or 'unfit'. During this case this was a binary classification problem (a yes no type problem).

2.1.4 RANDOM FOREST

Random Forest may be a popular machine learning algorithm that belongs to the supervised learning technique. It are often used for both Classification and Regression problems in ML. It is supported the concept of ensemble learning, which may be a process of mixing multiple classifiers to unravel a posh problem and to enhance the performance of the model.

As the name suggests, "Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset." rather than counting on one decision tree, the random forest takes the prediction from each tree and supported the bulk votes of predictions, and it predicts the final output.

The greater number of trees within the forest results in higher accuracy and prevents the matter of over fitting.

2.1.5 SVM

A type of learning algorithm developed in 1990 is the Support Vector Machine (SVM). This approach is based on observations from the theory of mathematical learning applied by VapNik.

SVM computers are also closely related to the functions of the kernel, which is a core principle for most learning tasks. In a number of fields, the kernel system and SVM are used. It involves the retrieval of multimedia data, bioinformatics, and pattern recognition.

2.2 UNSUPERVISED LEARNING

In unsupervised learning, we'd like to find out function f when only all the input examples are given and no corresponding output results are available. Unsupervised learning, in comparison to supervised form of learning, permits a user to learn models, which are complex and larger. Supervised learning is inefficient in learning models that contain deep hierarchies, because difficulty of learning the tasks increases exponentially in such cases. Whereas in unsupervised learning, the method of learning can proceed hierarchically from the observations into more abstract level of representation. The learning time during this case increases (approximately) linearly in terms of the amount of levels in model hierarchy because every additional hierarchy requires learning just one step. Furthermore, the underlying relationship between the input and output observations is difficult during this form, compared to supervised sort of learning. For example, clustering and discovery fall in unsupervised learning category. In clustering , a method is devised to form natural classes from a given dataset of examples or instances. In discovery, a selected goal isn't given; it's both inductive and deductive learning (explained later), during which the system learns with none help from an educator.

2.3 REINFORCEMENT LEARNING

Some action is performed on There is a midway between supervised and unsupervised learning, which is known as reinforcement learning. The environment by the training network during this method then, the training Based on the response obtained, the action is graded by the system as rewarding, i.e., good action, or punishing action, i.e., bad action. Network obtains a feedback reaction from it.

3. CONCLUSION

It is possible for machine learning to be supervised or unsupervised. If You have a smaller volume of details and clearly labeled data for education, select supervised Instruction. Non-supervised generally speaking, learning will deliver improved results and Results for broad sets of data. If you can easily get a massive data set, Accessible, go for techniques for deep learning. The processes You too, have Learned Reinforcement and Deep Reinforcement Learning Teaching. This paper surveys a number of different algorithms for machine learning. Each and every citizen today It consciously or unknowingly uses machine learning. From having a product that is recommended for online shopping to Image alerts on social networking websites. This essay offers an introduction to most of the common algorithms for machine learning.

REFERENCES

[1] https://www.spotlightmetal.com/machine- learning--definition-and-application-examples.

[2] W. Richert, L. P. Coelho, “Building Machine Learning Systems with Python”, Packt Publishing Ltd., ISBN 978-1-78216-140-0

[3] M. L. Marais, J. Patel, and M. Wolfson. The experimental design of classification models: An application of recursive partitioning and bootstrapping to commercial bank loan classifications. Journal of Accounting Research,22:87–114, 1984.

[4].https://www.xoriant.com/blog/product-engineering/decision-trees-machine-learning-algorithm

[5]. Ayon Dey, Machine Learning Algorithms: A Review, Vol. 7 (3) , 2016, 1174-1179,

[6]. Ali Heydarzadegan, Yaser Nemati, Mohsen Moradi1,” Evaluation of Machine Learning Algorithms in Artificial Intelligence”, IJCSMC, Vol. 4, Issue. 5, May 2015.