International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 06 | June -2020 ISSN: 2582-3930

Overview of Supervised Machine Learning Methods & Future Scope

Vinay patel
Sharda University, Greater Noida

Abstract

Supervised machine learning is the search for algorithms that reason from externally supplied instances to produce general hypotheses, which then make predictions about future instances. In other words, the goal of supervised learning is to build a concise model of the distribution of class labels in terms of predictor features. The resulting classifier is then used to assign class labels to the testing instances where the values of the predictor features are known, but the value of the class label is unknown. This paper describes various supervised machine learning classification techniques. Of course, a single article cannot be a complete review of allsupervised machine learning classification algorithms (also known induction classification algorithms), yet we hope that the references cited will cover the major theoretical issues, guiding the researcher in interesting research directions and suggesting possible bias combinations that have yet to be explored.

Key Words: Supervised Machine Learning, Supervised Machine Learning Algorithms Future Scope

Introduction

1.1 Introduction of Supervised Learning

Machine learning represents a large field presented in information technology, statistics, probability, artificial intelligence, psychology, neurobiology and many other disciplines. With machine learning the problems can be solved simply by building a model that is a good representation of a selected dataset. Machine learning has become an advanced area from teaching the computers to mimic the human brain, and has brought the field of statistic to a broad discipline that produces fundamental statistical computational theories of the learning processes.

Machine learning is all about creating algorithms that allow the computer to learn. Learning is a process of finding statistical regularities or other patterns of data. The machine learning algorithms are created to be able to represent the human approach of learning some task. These algorithms can also represent an insight into relative difficulty of learning in different environments.

These days, the development of new computing technologies in the area of Big Data, machine learning is not like machine learning was inthe past. Today, many of the machine learning algorithms have been developed, updated and improved and the recent development in machine learning becomes the ability to automatically apply a variety of complex mathematical calculation to a big data, which calculates the results much faster.

The adaptive programming is very popular. It is used in machine learning where the applications are capable to recognize patterns, learning from experience, abstract new information from data or optimize the accuracy and efficiency of its processing and output. Also, the machine learning techniques are used to work with multidimensional data which are present in diverse amount of application areas.

So, based on the desired outcome of the algorithm, the machine learning algorithms are organized in the following groups:

- **Supervised Learning:** - The various algorithms generate a function that maps inputs to desired outputs. One standard formulation of the supervised learning task is the classification problem: the learner is required to learn (to approximate the behavior of) a function which maps a vector into one of several classes by looking at several input-output examples of the function.

- **Unsupervised Learning:-** Unsupervised learning is a type of machine learning algorithm used to draw inferences from datasets consisting of input data without labeled responses. The most common unsupervised learning method is cluster analysis, which is used for exploratory data analysis to find hidden patterns or grouping in data. The clusters are modeled using a measure of similarity which is defined upon metrics such as Euclidean or probabilistic distance.
- **Semi-Supervised Learning: -** Semi-supervised learning is an approach to machine learning that combines a small amount of labeled data with a large amount of unlabeled data during training. Semi-supervised learning falls between unsupervised learning (with no labeled training data) and supervised learning (with only labeled training data).
- **Reinforcement Learning: -** Reinforcement learning (RL) is an area of machine learning concerned with how software agents ought to take actions in an environment in order to maximize the notion of cumulative reward. Reinforcement learning is one of three basic machine learning paradigms,alongside supervised learning and unsupervised learning.

Besides these groups of machine learning algorithms, they are basically divided into two general groups, supervised and unsupervised learning.

In supervised algorithms, the classes are predetermined. These classes are created in a manner of finite set, defined by the human, which in practice means that a certain segment of data will be labeled with these classifications. The task of the machine learning algorithm is to find patterns and construct mathematical models. These models are then evaluated based on the predictive capacity in relation to measures of variance in the data itself.

It is also useful to make difference between two main supervised models:

classification models (classifiers) and regression models. Regression models map the input space into a real-value domain. The classifiers map the input space into pre- defined classes. There are many alternatives for representing classifiers, for instance, support vector machines, decision trees, probabilistic summaries, algebraic function, etc. Along with regression and probability estimation, classification is one of the most studied models, possibly one with the greatest practical relevance. The potential benefits of progress in classification are immense since the technique has great impact on other areas, both within Data Mining and in its applications.

On the other hand, the unsupervised learning algorithms are not provided with classifications. The main task of unsupervised learning is to automatically develop classifications labels. These algorithms are searching the similarity between pieces of data in order to determinate if they can be categorized and create a group. These groups are so called clusters, and they represent whole family of clustering machine learning techniques. In this unsupervised classification (cluster analysis) the machine doesn't know how the clusters are grouped. Using the cluster analysis, there is a bigger potential for surprising ourselves. Thus, cluster analysis is a very promising tool for the exploration of relationships between many papers.

This paper is a representation of different types of supervised machine learning algorithms and their most efficient use to make decisions more efficient and to complete the task in more optimized form. In this paper, how different algorithms give the machine different learning experience and are adopting other things from the environment will be shown, and after which the machine makes a decision and performs specialized tasks.

The paper is organized as follows: Section II paper takes us into consideration the main related work that are used for completing this paper. Section III provides the overview of the supervised machine learning process. Section IV discusses the various learning algorithms used to perform learning process.

1.2 Block of Diagram of Concept (Supervised Learning)

The learning process in a simple machine learning model is divided into two steps: training and testing. In training process, samples in training data are taken as input in which features are learned by learning algorithm or learner and build the learning model. In the testing process, learning model uses the execution engine to make the prediction for the test or production data. Tagged data is the output of learning model which gives the final prediction or classified data.

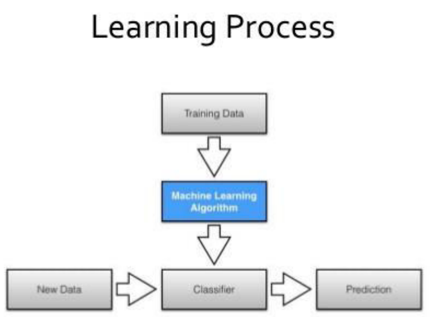


Figure 1: Supervised Learning Process

Supervised learning (Figure 1) is the most common technique in the classification problems, since the goal is often to get the machine to learn a classification system that we've created.

Most commonly, supervised learning leaves the probability for input undefined, such as an input where the expected output is known. This process provides dataset consisting of features and labels. The main task is to construct an estimator able to predict the label of an object given by the set of features. Then, the learning algorithm receives a set of features as inputs along with the correct outputs and it learns by comparing its actual output with corrected outputs to find errors. It then modifies the model accordingly. The model that is created is not needed as long as the inputs are available, but if some of the input values are missing, it is not possible to infer anything about the outputs.

Supervised learning is the most common technique for training for neutral networks and decision trees. Both of these are depended on the information given by the pre-determinate classification.

Also, this learning is used in applications where historical data predicts likely feature events. There are many practical examples of this learning, for instance an application that predicts the species of iris given a set of measurements of its flower. As previously mentioned, the supervised learning tasks are divided into two categories: classification and regression. In classification, the label is discrete, while
in regression, the label is continuous.

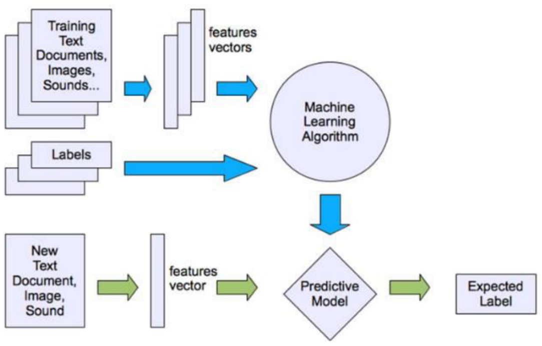


Figure 2: Supervised Learning Model

As shown on Figure 2, the algorithm makes the distinction between the observed data X that is the training data, in most cases structured data given to the model during the training process. In this process, the supervised learning algorithm builds the predictive model. After its training, the fitted model would try to predict the most likely labels for a new set of samples X in the testing set. Depending on the nature of the target y, supervised learning can be classified:

- If y has values in a fixed set of categorical outcomes (integers), the task to predict y is called classification
- If y has floating point values, the task to predict y is called regression

1.3 How it works of Supervised Learning

This algorithm consists of a target / outcome variable (or dependent variable) which is to be predicted from a given set of predictors (independent variables). Using these set of variables, we generate a function that map inputs to desired outputs. The training process continues until the model achieves a desired level of accuracy on the training data. Examples of Supervised Learning: Regression, Decision Tree, Random Forest, KNN, Logistic Regression etc.

1.4Challenges of Supervised Machine Learning

Here, are challenges in faced in supervised machine learning

- Irrelevant input feature present training data could give inaccurate results.

- Data preparation and pre-processing is always a challenge.
- Accuracy suffers when impossible, unlikely, and incomplete values have been inputted as training data.
- If the concerned expert is not available, then the other approach is "brute-force." It means you need to think that the right features (input variables) to train the machine on. It could be inaccurate.

1.5 Advantage of Supervised Learning:

There are many advantages of supervised machine learning:

- Supervised learning allows you to collect data or produce a data output from the previous experience.
- Helps you to optimize performance criteria using experience.
- Supervised machine learning helps you to solve various types of real-world computation problems.

1.6 Disadvantage of Supervised Learning

There are many disadvantages of supervised machine learning.

- Decision boundary might be overstrained if your training set which doesn't have examples that you want to have in a class
- You need to select lots of good examples from each class while you are training the classifier.
- Classifying big data can be a real challenge.
- Training for supervised learning needs a lot of computation time.

1.7 Best Practices of Supervised Machine Learning

- Before doing anything else, you need to decide what kind of data is to be used as a training set.
- You need to decide the structure of the learned function and learning algorithm.
- Gatherer corresponding outputs either from human experts or from measurements.

2. Supervised Learning Algorithms

2.1 Linear Regression

It represents a modeling relationship between a continuous scalar dependent variable y (also label or target in machine learning terminology) and one or more (a D-dimensional vector) explanatory variables (also independent variables, input variables, features, observed data, observations, attributes, dimensions, data point, etc.) denoted *X* using a linear function. In regression analysis the goal is to predict a continuous target variable, whereas another area called classification is predicting a label from a finite set. The model for a multiple regression which involves linear combination of input variables takes the form:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + e$$

Linear regression also belongs to the category of supervised learning algorithms. It means we train the model on a set of labeled data (training data) and then use the model to predict labels on unlabeled data (testing data).

As shown on Figure 4, the model (red line) is calculated using training data (blue points) where each point has a known label (*y* axis) to fit the points as accurately as possible by minimizing the value of a chosen loss function. We can then use the model to predict unknown labels (we only know *x* value and want to predict *y* value).

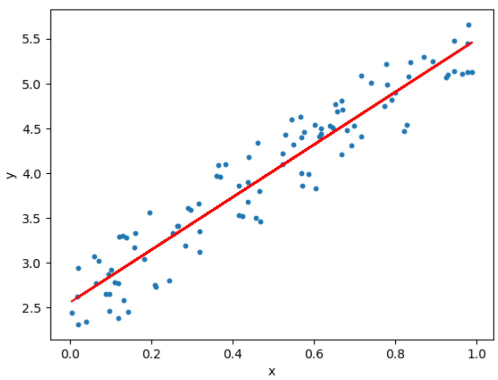


Figure 3: Visual representation of the linear regression

2.2 Decision Tree

Decision tree represents a classifier expressed as a recursive partition of the instance space. The decision tree consists of nodes that form so called root tree, which means that it is a distributed tree with a basic node called root with no incoming edges.

All of the other nodes have exactly one incoming edge. The node that has outgoing edges is called internal node or a test node. The rest of the nodes are called leaves. In a decision tree, each test node splits the instance space into two or more sub-spaces according to a certain discrete function of the input values. In the simplest case, each test considers a single attribute, such that the instance space is portioned according to the attribute's value. In case of numeric attributes, the condition refers to a range.

Each leaf is assigned to one class that represents the most appropriate target value. The leaf may hold a probability vector that indicates the probability of the target attribute having a certain value. The instances are classified by navigating them from the root of the tree down the leaf, according to the outcome of the tests along the path. On Figure 3 describes a simple use of the decision tree. Each node is labeled with the attribute it tests, and its branches are labeled with its corresponding values.

Given this classifier, the analyst can predict the response of some potential customer and understanding the behavioral characteristics of the entire potential customers' population

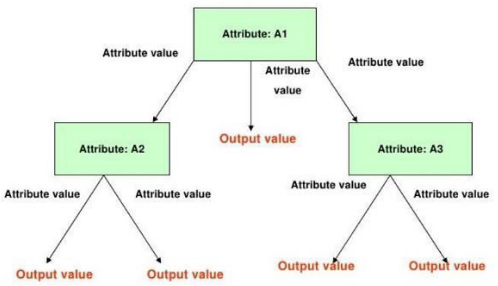


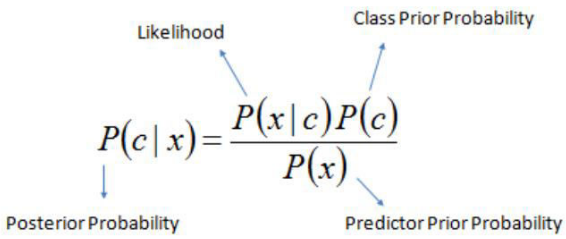
Figure 4: Decision Tree Example

1.3 Naive Bayes

It is a classification technique based on Bayes' theorem with an assumption of independence between predictors. In simple terms, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature. For example, a fruit may be considered to be an apple if it is red, round, and about 3 inches in diameter. Even if these features depend on each other or upon the existence of the other features, a naive Bayes classifier would consider all of these properties to independently contribute to the probability that this fruit is an apple.

Naive Bayesianmodel is easy to build and particularly useful for very large data sets. Along with simplicity, Naive Bayes is known to outperform even highly sophisticated classification methods.

Bayes theorem provides a way of calculating posterior probability P(c/x) from P(c), P(x) and P(x/c). Look at the equation below:



$$P(c \mid x) = \frac{P(x \mid c)P(c)}{P(x)}$$

$$P(c \mid X) = P(x_1 \mid c) \times P(x_2 \mid c) \times \cdots \times P(x_n \mid c) \times P(c)$$

Figure 5: Naïve Bayes

Here,

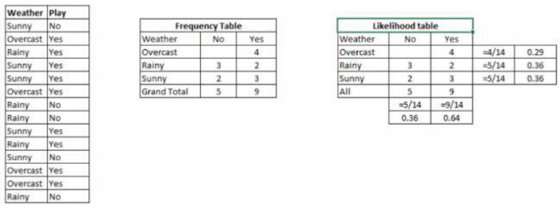
- P(c/x) is the posterior probability of class (target) given predictor (attribute).
- P(c) is the prior probability of class.
- P(x/c) is the likelihood which is the probability of predictor given class.
- P(x) is the prior probability of predictor.

Example: Let's understand it using an example. Below I have a training data set of weather and corresponding target variable 'Play'. Now, we need to classify whether players will play or not based on weather condition. Let's follow the below steps to perform it.

Step 1: Convert the data set to frequency table

Step 2: Create Likelihood table by finding the probabilities like Overcast probability = 0.29 and probability of playing is 0.64.

Volume: 04 Issue: 06 | June -2020 ISSN: 2582-3930



Weather	Play
Sunny	No
Overcast	Yes
Rainy	Yes
Sunny	Yes
Sunny	Yes
Overcast	Yes
Rainy	No
Rainy	No
Sunny	Yes
Rainy	Yes
Sunny	No
Overcast	Yes
Overcast	Yes
Rainy	No

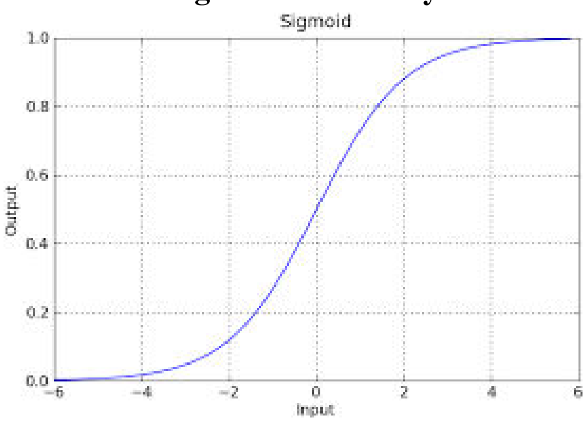
Frequency Table

Weather	No	Yes
Overcast		4
Rainy	3	2
Sunny	2	3
Grand Total	5	9

Likelihood table

Weather	No	Yes		
Overcast		4	=4/14	0.29
Rainy	3	2	=5/14	0.36
Sunny	2	3	=5/14	0.36
All	5	9		
	=5/14	=9/14		
	0.36	0.64		

Figure 6: Naïve Bayes



Step 3: Now, use Naive Bayesian equation to calculate the posterior probability for each class. The class with the highest posterior probability is the outcome of prediction.

Problem: Players will pay if weather is sunny, is this statement is correct?

We can solve it using above discussed method, so P (Yes | Sunny) = P (Sunny | Yes) * P(Yes) / P (Sunny)

Here we have P (Sunny |Yes) = 3/9 = 0.33, P(Sunny) = 5/14 = 0.36, P(Yes)= 9/14 = 0.64

Now, P (Yes | Sunny) = 0.33 * 0.64 / 0.36 = 0.60, which has higher probability.

Naive Bayes uses a similar method to predict the probability of different class based on various attributes. This algorithm is mostly used in text classification and with problems having multiple classes.

2. 4 Logistic Regression

Like the naive Bayes, logistic regression works by extracting some set of weighted features from the input, taking logs and combining them linearly, which means that each feature is multiplied by a weight and then added up.

The most important difference between naive Bayes and logistic regression is that the logistic regression is a discriminative classifier while the naive Bayes is a generative classifier.

Logistic regression is a type of regression that predicts the probability of occurrence of an event by fitting data to a logistic function. Just as many forms of regression analysis, logistic regression makes use of several predictor variables that may be numerical or categorical.

The logistic regression hypothesis is defined as:

$$h_{\theta}(x) = g(\theta^{T}x)$$

Where the function $g$is sigmoid function defined as:

$$g(z) = \frac{1/}{1 + e^{-z}}$$

The sigmoid function has special properties that result the values in range [0,1], as visualized on Figure 7

Figure 7: Visual representationof the Logistic Function

$J(\theta) = 1\ m\sum[-y\ (i) \log\ (h\theta(x\ (i)\ ))\ -(1 - y\ (i)\ )\ \log(1 - h\theta(x\ (i)\ ))]$

To find the minimum of this cost function in machine learning we will use a built-in function called fmin bfgs2, which finds the best parameters θ for the logistic regression cost function given a fixed dataset (of x and y values). The parameter are the initial values of the parameter that need to be optimized and a function that when the traning set and a particular θ, computes the logistic regression cost and gradient with respect to θ for the dataset woth x and y values. The final θ value will be used to plot the decision boundary of the traning data.

Summary

- In Supervised learning, you train the machine using data which is well "labelled."
- You want to train a machine which helps you predict how long it will take you to drive home from your workplace is an example of supervised learning.
- Regression and Classification are two types of supervised machine learning techniques.

- Supervised learning is a simpler method while Unsupervised learning is a complex method.
- The biggest challenge in supervised learning is that Irrelevant input feature present training data could give inaccurate results.
- The main advantage of supervised learning is that it allows you to collect data or produce a data output from the previous experience.
- The drawback of this model is that decision boundary might be overstrained if your training set doesn't have examples that you want to have in a class.
- As a best practice of supervise learning, you first need to decide what kind of data should be used as a training set.

Machine Learning Future Scope

Machine Learning can be a competitive advantage to any company be it a top MNC or a startup as things that are currently being done manually will be done tomorrow by machines. Machine Learning revolution will stay with us for long and so will be the future of Machine Learning.

Machine Learning is one of the best career choices of the 21st century. It has plenty of job opportunities with a high-paying salary. Also, the future scope of Machine Learning is on its way to make a drastic change in the world of automation. Further, there is a wide scope of Machine Learning in India. Thus, you can make a lucrative career in the field of Machine Learning to contribute to this growing digital world. In this blog, we will discuss various trends and the future scope of Machine Learning.

The scope of Machine Learning is not limited to the investment sector. Rather, it is expanding across all fields such as banking and finance, information technology, media & entertainment, gaming, and the automotive industry. As the Machine Learning scope is very high, there are some of the areas where researchers are working toward revolutionizing the world for the future. Like Automotive industry, Robotics, Quantum Computing, Computer Vision.

Machine Learning Job Scope and Salary Trends: -

The scope of Machine Learning in India, as well as in other parts of the world, is high in comparison to other career fields when it comes to job opportunities. According to Gartner, there will be 2.3 million jobs in the field of Artificial Intelligence and Machine Learning by 2022. Also, the salary of a Machine Learning Engineer is much higher than the salaries offered to other job profiles.

According to Forbes, the average salary of a Machine Learning Engineer in the United States is US$99,007. In India, it is ₹865,257. Let us look at the graph of top job profiles listed by Indeed.

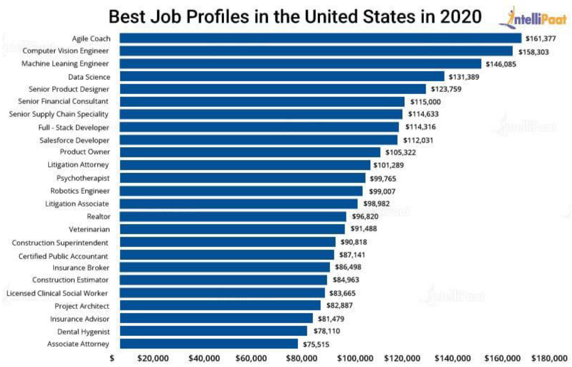


Figure 8: Future Scope of Machine Learning

This shows that the Machine Learning scope is extremely high in terms of salary and the Proption to make a lucrative career in ML by becoming a Machine Learning professional. Further, in this blog on the future scope of Machine Learning, we will look into the skills that are required to become an ML Engineer

Conclusion

As discussed in the paper, for the supervised learning it may be concluded that is one of the dominant methodologies in machine learning. The techniques that are used are even more successful than the unsupervised techniques because the ability of labelled training data provide us clearer criteria for model optimization. The supervised learning methods contain a large set of algorithms which are improving all the time by the data scientists.

This paper provides an overview of couple of supervised learning algorithms. There is a brief explanation of the machine learning process. This paper also describes the basic structure of some various machine learning algorithms and their basic structure.

This area has the attention from many developers and has gained substantial progress in the last decade. The learning methods achieved excellent performance that would have been difficult to obtain in the previous decades. Because of the rapid progression, there is plenty of space for the developers to work or to improve the supervised learning methods and their algorithms.

Reference

[1] Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 2nd Edition https://learning.oreilly.com/library/view/hands-on-machine-learning/9781492032632/

[2] Anish Talwar, Yogesh Kumar. (2013). Machine Learning: An artificial intelligence methodology. In International Journal of Engineering and Computer Science.

[3] https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/

[4] https://www.guru99.com/supervised-machine-learning.html

[5] https://www.geeksforgeeks.org/ml-types-learning-supervised-learning/

Author: **vinay patel,** I am pursuing MCA (Master of Computer Application) from Sharda University, Greater Noida, My Research is Supervised Machine Learning Methods & Future Scope