ISSN NO: 2770-2367
Date of Publication: 28-04-2022

The Importance of Supervised Learning in the Operation of Modern Techniques

Babajanov Boburbek Farhodovich
Student, Urgench branch of TUIT, Email: babajonovboburbek@gmail.com
Ollaberganova Muyassar Davlatboy qizi
Student, Urgench branch of TUIT, Email: ollaberganovamuyassar@gmail.com
Ravshanqulov Abbos Ravshanqul o'g'li
Student, Urgench branch of TUIT, Email: ravshanqulovabbo3@gmail.com
Abdalova Hilola Batirovna
Student, Urgench branch of TUIT, Email: xabdalova@gmail.com

Abstract. This article discusses the importance of supervised learning and how it is widely used today. In addition, information was provided on the types of Supervised Learning, such as Classification and Regression, and the advantages and disadvantages of this learning. In addition, the importance of this training in the work of modern technologies is explained.

Key words. Machine Learning, Classification, Regression, predictor, prediction, algorithm, spam, model.

Supervised learning, also known as supervised machine learning, is a subcategory of machine learning and artificial intelligence. It is defined by its use of labeled datasets to train algorithms that to classify data or predict outcomes accurately. As input data is fed into the model, it adjusts its weights until the model has been fitted appropriately, which occurs as part of the cross validation process. Supervised learning helps organizations solve for a variety of real-world problems at scale, such as classifying spam in a separate folder from your inbox.

Supervised learning uses a training set to teach models to yield the desired output. This training dataset includes inputs and correct outputs, which allow the model to learn over time. The algorithm measures its accuracy through the loss function, adjusting until the error has been sufficiently minimized.

Supervised learning can be separated into two types of problems when data mining—classification and regression:

1. **Classification** uses an algorithm to accurately assign test data into specific categories. It recognizes specific entities within the dataset and attempts to draw some conclusions on how those entities should be labeled or defined. Common classification algorithms are linear classifiers, support vector machines (SVM), decision trees, k-nearest neighbor, and random forest, which are described in more detail below.
2. **Regression** is used to understand the relationship between dependent and independent variables. It is commonly used to make projections, such as for sales revenue for a given business. Linear regression, logistical regression, and polynomial regression are popular regression algorithms.

What is the Classification Algorithm?

The Classification algorithm is a Supervised Learning technique that is used to identify the category of new observations on the basis of training data. In Classification, a program learns from the given dataset or observations and then classifies new observation into a number of classes or groups. Such as, Yes or No, 0 or 1, Spam or Not Spam, cat or dog, etc. Classes can be called as targets/labels or categories.

ISSN NO: 2770-2367
Date of Publication: 28-04-2022

Unlike regression, the output variable of Classification is a category, not a value, such as "Green or Blue", "fruit or animal", etc. Since the Classification algorithm is a Supervised learning technique, hence it takes labeled input data, which means it contains input with the corresponding output.

In classification algorithm, a discrete output function(y) is mapped to input variable(x).

y=f(x), where y = categorical output

The best example of an ML classification algorithm is **Email Spam Detector**.

What is the Regression Algorithm?

Regression analysis is a statistical method to model the relationship between a dependent (target) and independent (predictor) variables with one or more independent variables. More specifically, Regression analysis helps us to understand how the value of the dependent variable is changing corresponding to an independent variable when other independent variables are held fixed. It predicts continuous/real values such as temperature, age, salary, price, etc.

In Regression, we plot a graph between the variables which best fits the given datapoints, using this plot, the machine learning model can make predictions about the data. In simple words, "Regression shows a line or curve that passes through all the datapoints on target-predictor graph in such a way that the vertical distance between the datapoints and the regression line is minimum." The distance between datapoints and line tells whether a model has captured a strong relationship or not.

Supervised learning is a process of providing input data as well as correct output data to the machine learning model. The aim of a Supervised learning algorithm is to find a mapping function to map the input **variable(x)** with the output **variable(y)**.

Steps Involved in Supervised Learning:

a) First Determine the type of training dataset

b) Collect/Gather the labelled training data.

c) Split the training dataset into training dataset, test dataset, and validation dataset.

d) Determine the input features of the training dataset, which should have enough knowledge so that the model can accurately predict the output.

e) Determine the suitable algorithm for the model, such as support vector machine, decision tree, etc.

f) Execute the algorithm on the training dataset. Sometimes we need validation sets as the control parameters, which are the subset of training datasets.

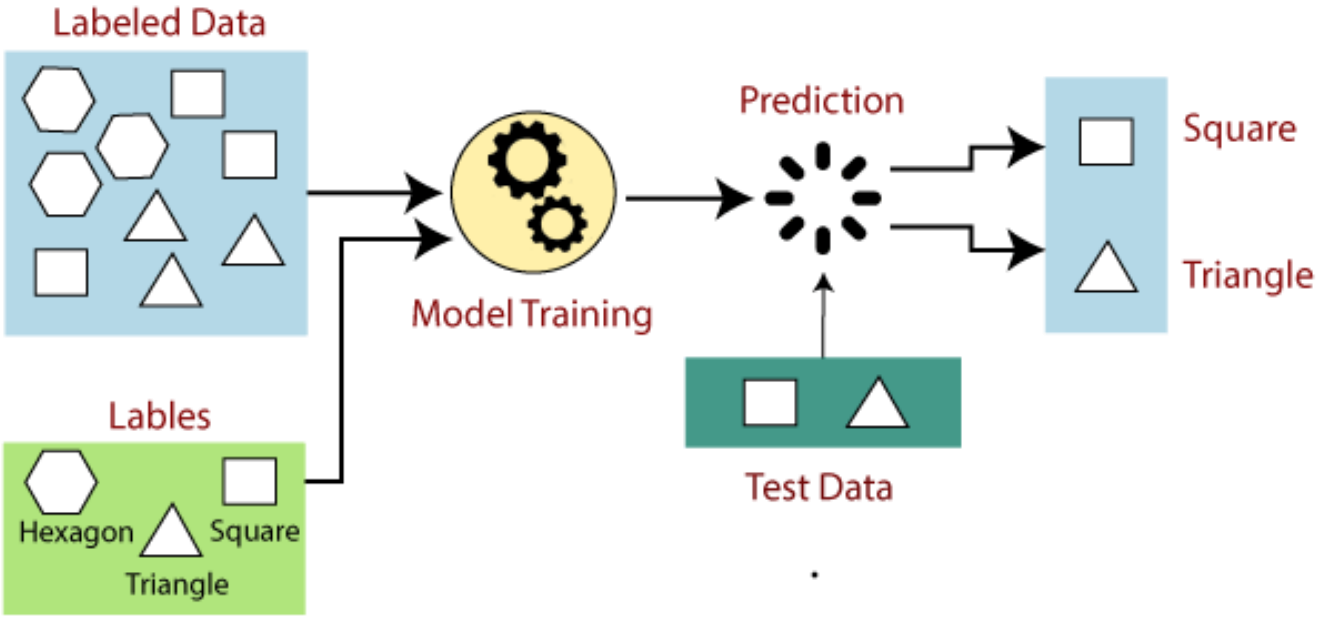
g) Evaluate the accuracy of the model by providing the test set. If the model predicts the correct output, which means our model is accurate.

In the real-world, Supervised learning can be used for Risk Assessment, Image classification, Fraud Detection, spam filtering, etc.

In Supervised learning, models are trained using labelled dataset, where the model learns about each type of data. Once the training process is completed, the model is tested on the basis of test data (a subset of the training set), and then it predicts the output.

The working of Supervised learning can be easily understood by the below example and diagram:

ISSN NO: 2770-2367
Date of Publication: 28-04-2022



Advantages of Supervised learning:

a) With the help of Supervised learning, the model can predict the output on the basis of prior experiences.
b) In Supervised learning, we can have an exact idea about the classes of objects.
c) Supervised learning model helps us to solve various real-world problems such as fraud detection, spam filtering, etc.

Disadvantages of supervised learning:

a) Supervised learning models are not suitable for handling the complex tasks.
b) Supervised learning cannot predict the correct output if the test data is different from the training dataset.
c) Training required lots of computation times.
d) In Supervised learning, we need enough knowledge about the classes of object.

Like all machine learning algorithms, supervised learning is based on training. During its training phase, the system is fed with labeled data sets, which instruct the system what output is related to each specific input value. The trained model is then presented with test data: This is data that has been labeled, but the labels have not been revealed to the algorithm. The aim of the testing data is to measure how accurately the algorithm will perform on unlabeled data.

The conclusion is that Supervised learning is now almost universal and widely used in many areas. These areas are medicine, education, police, bioengineering, chemical technology and others. The development of such technologies is one of the systems that can greatly facilitate many difficult tasks of people.

References

1. Bessmertny I.A. Artificial intelligence - St. Petersburg: St. Petersburg State University ITMO, 2010. - 132 p.
2. Brink H., Richards D., Feverolf M. Machine learning. - St. Petersburg: Peter, 2017. -336 p.: ill. – (Series "Programmer's Library").
3. Brockman D. What do we think about machines that think: Leading world artificial intelligence scientists. M.: - Alpina non-fiction, 2017. - 552 p.

ISSN NO: 2770-2367
Date of Publication: 28-04-2022

4. China spent $279 billion on research and development last year USD – Access mode: https://hightech.fm/2018/02/27/279-billion-on-rd.
5. Kretsu S. Artificial intelligence in business - the experience of Russian brands. Access mode: https://vc.ru/25645-ai-business
6. Osipov G. Artificial Intelligence: The state of research and a look at the future. – Access mode: http://www.raai.org/about/persons/osipov/pages/ai/ai.html
7. Demchenko D. Map of the application of artificial intelligence technologies: Medicine, education, transport and other areas. – Access mode: https://vc.ru/p/ai-map.