

Aim- Perform Classification Analysis for the numerical classes (0 to 9).

Objectives- Implementation of Support Vector Machines (SVM) for classifying images of handwritten digits into their respective numerical classes (0 to 9).

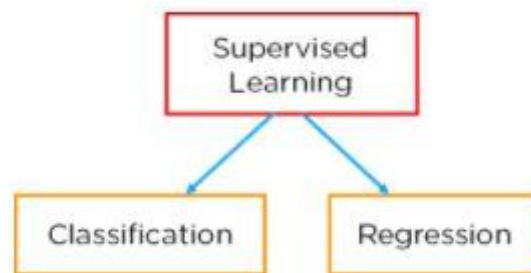
Theory-

In Supervised Learning, the model learns by example. Along with our input variable, we also give our model the corresponding correct labels. While training, the model gets to look at which label corresponds to our data and hence can find patterns between our data and those labels.

Some examples of Supervised Learning include:

1. It classifies spam Detection by teaching a model of what mail is spam and not spam.
2. Speech recognition where you teach a machine to recognize your voice.
3. Object Recognition by showing a machine what an object looks like and having it pick that object from among other objects.

We can further divide Supervised Learning into the following:



What is Classification?

Classification is defined as the process of recognition, understanding, and grouping of objects and ideas into preset categories, "sub-populations." With the help of these pre-categorized training datasets, classification in machine learning programs leverage a wide range of algorithms to classify future datasets into respective and relevant categories.

Classification algorithms used in machine learning utilize input training data for the purpose of predicting the likelihood or probability that the data that follows will fall into one of the

predetermined categories. One of the most common applications of classification is for filtering emails into “spam” or “non-spam”, as used by today’s top email service providers.

In short, classification is a form of “pattern recognition.” Here, classification algorithms applied to the training data find the same pattern (similar number sequences, words or sentiments, and the like) in future data sets.

What is Classification Algorithm?

Based on training data, the Classification algorithm is a Supervised Learning technique used to categorize new observations. In classification, a program uses the dataset or observations provided to learn how to categorize new observations into various classes or groups. For instance, 0 or 1, red or blue, yes or no, spam or not spam, etc. Targets, labels, or categories can all be used to describe classes. The Classification algorithm uses labeled input data because it is a supervised learning technique and comprises input and output information.

In simple words, classification is a type of pattern recognition in which classification algorithms are performed on training data to discover the same pattern in new data sets.

Classification Predictive Modeling

A classification problem in machine learning is one in which a class label is anticipated for a specific example of input data.

Problems with categorization include the following:

- Give an example and indicate whether it is spam or not.
- Identify a handwritten character as one of the recognized characters.
- Determine whether to label the current user behavior as churn.

A training dataset with numerous examples of inputs and outputs is necessary for classification from a modeling standpoint.

- A model will determine the optimal way to map samples of input data to certain class labels using the training dataset. The training dataset must therefore contain a large number of samples of each class label and be suitably representative of the problem.

- When providing class labels to a modeling algorithm, string values like "spam" or "not spam" must first be converted to numeric values. Label encoding, which is frequently used, assigns a distinct integer to every class label, such as "spam" = 0, "no spam," = 1.
- There are numerous varieties of algorithms for classification in modeling problems, including predictive modeling and classification.
- It is typically advised that a practitioner undertake controlled tests to determine what algorithm and algorithm configuration produces the greatest performance for a certain classification task because there is no strong theory on how to map algorithms onto issue types.
- Based on their output, classification predictive modeling algorithms are assessed. A common statistic for assessing a model's performance based on projected class labels is classification accuracy. Although not perfect, classification accuracy is a reasonable place to start for many classification jobs.
- Some tasks may call for a class membership probability prediction for each example rather than class labels. This adds more uncertainty to the prediction, which a user or application can subsequently interpret. The ROC Curve is a well-liked diagnostic for assessing anticipated probabilities.

There are four different types of Classification Tasks in Machine Learning and they are following -

- Binary Classification
- Multi-Class Classification
- Multi-Label Classification
- Imbalanced Classification

Conclusion- We implements Support Vector Machines (SVM) for classifying images of handwritten digits into their respective numerical classes (0 to 9).