

PRIOR KNOWLEDGE

Team ID : PNT2022TMID02158

Project Name : A novel method for handwritten digit

1. SUPERVISED AND UNSUPERVISED LEARNING

Supervised learning, as the name indicates, has the presence of a supervisor as a teacher. Basically supervised learning is when we teach or train the machine using data that is well labelled. Which means some data is already tagged with the correct answer. After that, the machine is provided with a new set of examples(data) so that the supervised learning algorithm analyses the training data(set of training examples) and produces a correct outcome from labelled data.

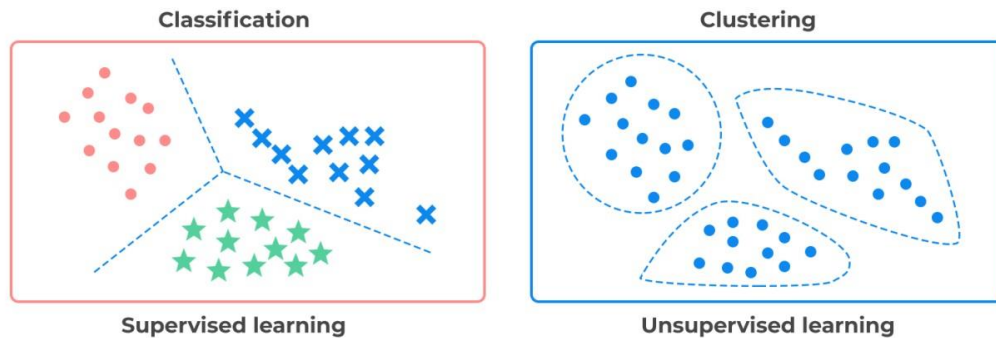
Types:-

- Regression
- Logistic Regression
- Classification
- Naive Bayes Classifiers
- K-NN (k nearest neighbors)
- Decision Trees
- Support Vector Machine

Unsupervised learning is the training of a machine using information that is neither classified nor labeled and allowing the algorithm to act on that information without guidance. Here the task of the machine is to group unsorted information according to similarities, patterns, and differences without any prior training of data.



Supervised vs. Unsupervised Learning



2. REGRESSION, CLASSIFICATION AND CLUSTERING

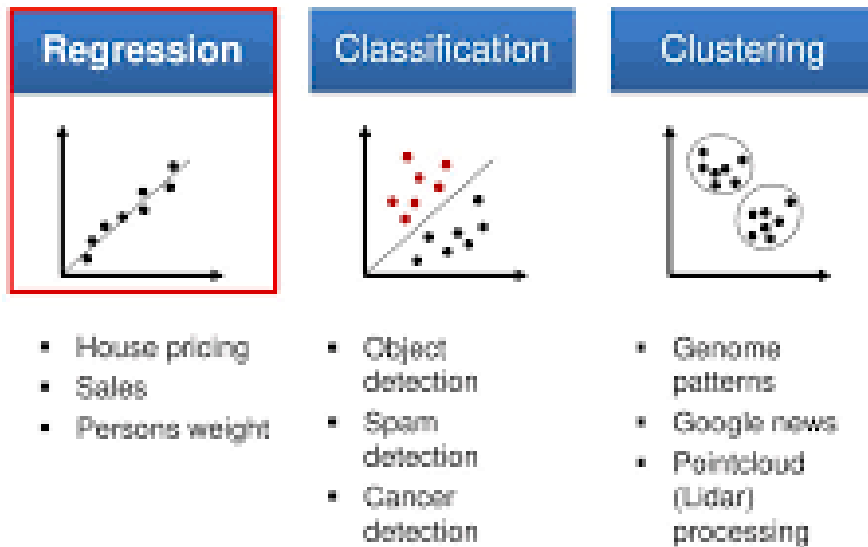
Regression and classification are supervised learning approach that maps an input to an output based on example input-output pairs, while clustering is a unsupervised learning approach.

Regression: It predicts continuous valued output. The Regression analysis is the statistical model which is used to predict the numeric data instead of labels. It can also identify the distribution trends based on the available data or historic data. Predicting a person's income from their age, education is example of regression task.

Classification: It predicts discrete number of values. In classification the data is categorized under different labels according to some parameters and then the labels are predicted for the data. Classifying emails as either spam or not spam is example of classification problem.

Clustering: Clustering is the task of partitioning the dataset into groups, called clusters. The goal is to split up the data in such a way that points within single cluster are very similar and points in different clusters are different. It determines grouping among unlabelled data.

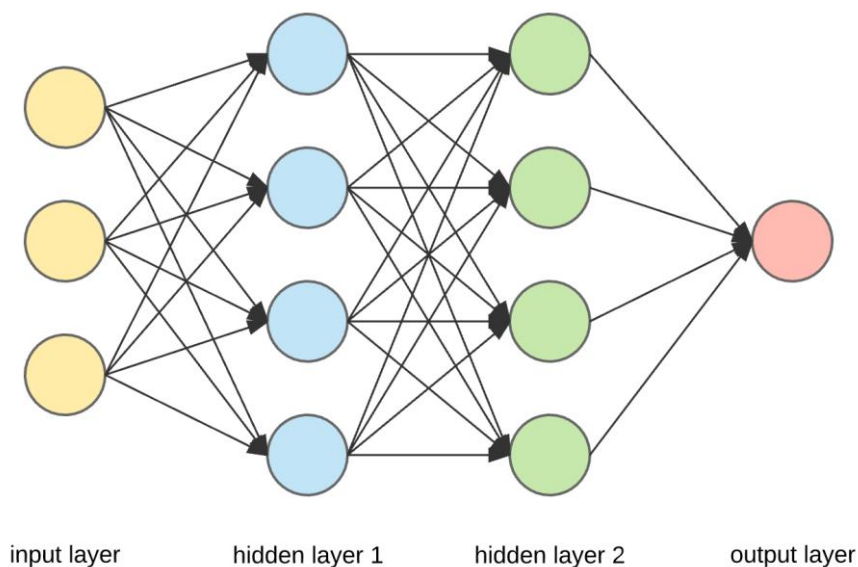
Motivation – Algorithms in Machine Learning



3 - 6

3. ARTIFICIAL NEURAL NETWORK

Artificial neural networks, usually simply called neural networks or neural nets, are computing systems inspired by the biological neural networks that constitute animal brains. An ANN is based on a collection of connected units or nodes called artificial neurons, which loosely model the neurons in a biological brain.



4. CONVOLUTIONAL NEURAL NETWORK

CNN is a Deep Learning algorithm that can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image, and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics.

