Unde Suvarna, 2023, 11:6
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology

An Open Access Journal

Machine learning techniques and algorithms: A survey

Unde Suvarna
Vishwabharati college of Engineering, Ahmednagar

Abstract- Machine learning, which analyzes research and creates algorithms based on prediction from data. It creates a model based on the material used to make decisions or predictions. Machine learning algorithms help bridge the perception gap. In these articles we learn many machine learning methods and techniques and explains the concepts and process of machine learning, which is a well-known branch of computer science.

Keywords- Machine learning, Supervised learning, unsupervised learning, Semi-supervised learning, Reinforcement Learning.

I. INTRODUCTION

Based on knowledge structures and skills, machine learning is a subfield of computer science. Machine learning is about statistical calculations and is specifically applied to predictions. Current machine learning research mainly focuses on natural language processing, computer vision, pattern recognition, computational knowledge, and knowledge representation. Machine learning techniques can be called modeling techniques when used in a business environment.

The field of machine learning has made significant progress over the last two decades. This field has become way to develop software suitable for computer vision, speech recognition, natural language processing, robot control and other applications.

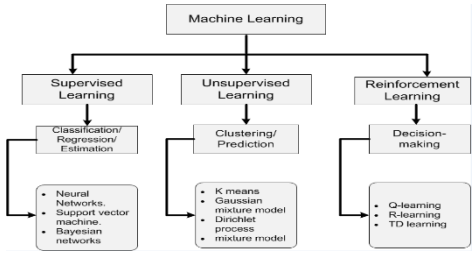

Fig.1 Types of ML

Types of Machine Learning

1. Supervised Machine Learning

Supervised learning is the most common type of machine learning. This way the model is trained on a list of data. In other words, the text that the model is trying to predict is added to the data. This model learns the mapping of features & labels during training. After training, the model can predict the consequences of new, unseen objects.

2. Unsupervised Machine Learning

Unsupervised learning involves training the model on unlabeled data. Models see patterns and relationships in their data. This type of learning is often used for merging and dimensionality reduction. While integration involves grouping similar data together, reducing dimensionality will reduce the number of variables required to obtain the key set.

3. Reinforcement Learning

Reinforcement learning is a type of machine learning in which an agent learns to make decisions by interacting with its environment. Agents are rewarded or punished based on the work they do to obtain the full reward. However, unlike supervised learning and unsupervised learning, complementary learning is especially suitable for problems that involve continuous information, and the decision of each step will affect the future.

II. ML ALGORITHMS

1. Decision Trees

A decision tree (DT) is a tree that classifies events by identifying them according to values; where each part of the decision tree represents a unique point in the sample that needs to be separated, and each branch represents what can be considered in terms of data value.

The examples are sort according to their main features, starting from the root node.

2. K- Means Clustering

The main idea of this algorithm is to divide N observations in space into K groups. The most up-to date Information and content are included in the category and serve as a template for the category. Therefore, the source data is divided into Voronoi units. The k-means algorithm is an iterative process that starts with selecting k words v1, v2…vk. In each iteration, the data points are divided into k groups based on the point closest to each point and the average is adjusted based on the points in the group. Providing information about the content according to the group average and updating the group definition according to the determined content continues until there is no change in the group average or content. Variants of K-means are called K-medoids. In k medoids, instead of averaging the cluster size, the information at the center is examined as an indicator of the cluster of interest.

3. Support Vector Machines (SVM)

SVM rotates around edges on either side of the plane separating two classes. To reduce the upper bound of the overall error, the main idea is to create the maximum distance between events on both sides and separate the hyper plane. Finding the optimal hyper plane is one of the fundamental concepts of linearly separable objects.

Data points at the edge of the ideal hyper plane are called support vector points, characterized by lines connecting these points. Other data points are ignored. The different features present in the training material do not affect the complexity of the SVM.

4. Neural Network

The neural network conceptual model was proposed in 1943. It consists of different cells. This unit receives information from other units, processes the feedback and sends it to other units. Since then, the development of neural network devices has been actively researched. Perceptron [16] is a neural network consisting of a single neuron that receives multiple inputs to generate an output. To identify discrete objects, a perceptron is used to find an m dimensional hyper plane in a given space Separating instances of two classes.

5. Q-learning

It is reinforcement learning is type of ML. It can also be called asynchronous dynamic programming (DP) method. Q learning allows operators to learn to mod on a Markov collection by analyzing the effects of their actions, and they do not need to create graphs. For the finite Markov decision process, Q-learning finds a good rule, provides an infinite and quasi stochastic search time, and improves the Predictive value of the total reward starting from the instantaneous state. . The optimal selection strategy may be related to Q-learning.

6. Random Forest Algorithm

Random Forest is a supervised learning algorithm that can be used in classification and regression problems in machine learning. It is an ensemble learning technique that combines multiple classifications to provide predictions and improve model performance. There are many decision trees that provide a range of data sets and visual averages to improve the model's predictive accuracy. A random forest should have 64 128 trees. The more trees there are, the more accurate the algorithm.

III. CONCLUSION

Various machine learning methods and methods are examined in this study. The main focus of this study is the classification of machine learning methods such as supervised learning, unsupervised learning, semi supervised learning and Reinforcement learning, and their various algorithms. We plan to

build a model based on machine learning in the future. Machine learning experts will be ready to learn more and will have a better learning strategy that can lead to better performance across the country.

REFERENCES

1. Cortes, Corinna, Vapnik, Vladimir N, "Support-vector networks". Machine Learning. 20 (3): 273–297, 1995.
2. Fisher, R. A, "The Use of Multiple Measurements in Taxonomic Problems". Annals of Eugenics. 7 (2): 179–188, 1936.
3. John C. Platt, "Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods". pp: 61-74, Advances in large margin classifiers MIT press, 1999.
4. Kaufman, L. and Rousseeuw, P.J, "Clustering by means of Medoids", in Statistical Data Analysis Based on the Norm and Related Methods, pp: 405–416, North-Holland, 1987.
5. Mac Queen J. B, "Some Methods for classification and Analysis of Multivariate Observations". Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability. pp: 281–297, University of California Press, 1967.
6. Duda, R. O. and Hart, P. E, "Pattern Classification and Scene Analysis". John Wiley and Sons, Inc, 1973.
7. Dempster, A.P, Laird N.M, Rubin, D.B, "Maximum Likelihood from Incomplete Data via the EM Algorithm". pp: 1–38, Journal of the Royal Statistical Society, 1977.
8. Baum, L. E.; Petrie, T, "Statistical Inference for Probabilistic Functions of Finite State Markov Chains". pp: 1554–1563, The Annals of Mathematical Statistics, 1966.
9. Ng, Andrew and Jordan, Michael, "On Discriminative vs. Generative classifiers:A comparison of logistic regression
10. Pearson, K., "On Lines and Planes of Closest Fit to Systems of Points in Space". pp: 559–572, Philosophical Magazine, 1901.
11. A. Carlson, J. Betteridge, B.Kisiel, B.Settles, E. R. Hruschka Jr, and T. M. Mitchell, Toward an architecture for never-ending language learning, AAAI, volume 5, 3, 2010
12. X. Chen, A. Shrivastava, and A. Gupta, Neil: Extracting visual knowledge from web data, In Proceedings of ICCV, 2013.