

The Machine Learning: The method of Artificial Intelligence

Ashis Kumar Ratha¹, Nisha Agrawal², Amisha Ananya Sikandar³

¹Ast.Prof, Department of Computer Science & Engg, VIT, Bargarh, Odisha, INDIA

^{2,3}Student Researcher, Department of Computer Science & Engg, VIT, Bargarh Odisha, INDIA

Abstract: In classification problem learning and decision making is at the core level of argument as well as artificial aspects. So scientists introduce machine learning which is widely used in artificial intelligence. Artificial intelligence planning systems have become an important tool for automating a wide variety of tasks. Machine Learning techniques enable a planning system to automatically acquired search control knowledge for different applications. In the field of robotics machine learning plays an important role, it helps in taking decision and increase the efficiency of the machine. Machine learning is used in much application which is the principle concept for intelligence system which assist to the ingenious introduction and advanced concepts of artificial intelligence.

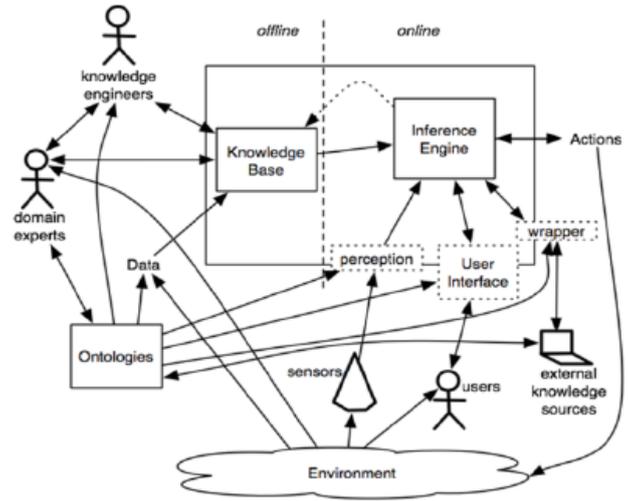


Fig1 : The Machine Learning Mechanism

1. INTRODUCTION

A type of artificial intelligence, which allows software applications whose accuracy in predicting output without explicitly programmed, is machine learning. Data mining and predictive modeling carries same process as machine learning. In this study of biological and artificial vision, learning used as a key.

To comprehend the virtual environment relating to the machine understanding different algorithm are introduced for avoiding to build the heavy machine having explicit programming. For taking independent decision different algorithm are implemented in different machines. A huge number of data sets being given to classify and based on these data sets it do some processing and tries to predict the result. Pattern recognition is the innate process of matching information from the environment with information store in memory. Pattern recognition is closely related to top down perception. In both cases knowledge and expectation are use in interpret information.

Pattern recognition involves detection of repeat characteristic, occurrences or some other attribute and this basic way to make sense of the word. In other way the constant attempt for identifying environmental information that matches the internal information.

In other words a branch of machine learning which used by many algorithms for getting optimized decisions is pattern recognition.

RELATED WORKS:

Sally Goldman et.al[1] gave the practical learning scenarios in which small amount of labeled data with a huge unlabeled data presented a co training strategy for using unlabeled data for improving the standard supervised learning algorithm. According to her assumption there are two types of hypothesis which defines the partitioning of instance space. Eg: The instance space with one equivalent class defined per tree in decision tree partition. The conclusion she gave was two supervised learning algorithm can be used successfully label data for each other.

Zoubin Ghahramani et.al[2] The overview of unsupervised learning from statistical modeling is provided by him. He concludes that unsupervised learning can be motivated from information theoretic and Bayesian principles. He further wind up that statistics provide a coherent framework for learning from data and reasoning and he divulge the types of model like graphical model which play a vital role in learning for handling of different kinds of data.

Rich Caruana et.al[3] The comparison between ten supervised learning methods has studied by him in supervised learning method introduced in last decade. These methods includes SVMs, neural nets, logistic regression, naïve bayes, memory based learning, random forest, decision trees, bagged trees, boosted tree and boosted stumps. To evaluate the learning methods they studied and examine the effect that calibrating the models through plat scaling and isotonic regression.

Niklas lavesson et.al[4] According to him performance is often only measured in terms of accuracy. Through cross validation tests, however some researches have given a different approach for evaluation of supervised learning. i.e measure function a limitation of current measure functions is that they can only handle two dimensional instance spaces. They present design and implementation of a generalized multidimensional measure function and demonstrate its use through a set of experiments. The result indicate that there are cases for which measure functions may be able to capture aspects of performance that cannot be captured by cross validation tests .Final result will be, they investigate the impact of learning algorithm parameter tuning.

Yogowati Praharsi et.al[5] The three supervised learning methods as k-nearest neighbor(k-NN),support vector data description(SVDD) and support vector machine(SVM) are approached by him because they do not suffer from the complexity of introducing a new class ,and further used for data description and classification. The output show that feature selection based on mean information gain and a standard deviation threshold can be considered as a substitute for forward selection.

PROBLEM FACED IN LEARNING:

As so many decisions are made, learning considered as a complex process depending upon machine to machine as well as algorithm to algorithm. From understanding a problem to responding, so many issues make a complex situation to respond for a machine, so it affects the learning process.

Perception defines how the machine perceives, so machine should also aim different types of challenges and environment to face. Though different inputs result different output, machine should considered only the optimized and appropriate output.

Problems faced during learning process are as follows:

Bias: Any error occur in learning algorithm is termed as bias. The problem faced during simultaneously minimizing two sources of error which prevent algorithm of supervised learning.

Noise: The unwanted data and the imperfection of data are now common in real world situation. The noise exist in the data are degrade the learning process but one of the properties of learning algorithm is to handling noisy data in all form.

Pattern Recognition: The next problem termed as pattern recognition, which aim is to providing reasonable answer of all inputs and performs the matching operation for all inputs and performs the matching operation for all inputs according to their statistical variation.

As machine is well known of mathematical models (square, rectangle, circle, etc), but it is also true that It become different for machine to process those inputs having different values.

Both the inputs and outputs are perceived in supervised learning .For responding all the inputs the algorithm has to generate all the inputs the algorithm has to generate all the training data from supervised learning.

When any agent is given immediate feedback supervised learning of action occurs. For solving any given problem using supervised learning, some steps to be carried out are:

1. Determination of training example and its type
2. Collecting training set
3. Knowledge of input feature of learned function
4. Determination of structure of learning function
5. Completion of decision, to run learning algorithm based on gathered set of data
- 6.Optimising the accuracy of learned function and performance of learning function and also performance should again measured on the set which are different from training set .

Algorithm	Predictive Accuracy	Fitting Speed	Prediction Speed	Memory Usage	Easy to Interpret	Handles categorical Predictors
Trees	Low	Fast	Fast	Low	yes	yes
SVM	High	Medium	*	*	*	No
Naive Bayes	Low	**	**	**	Yes	Yes
Nearest neighbour	***	Fast***	Medium	High	No	Yes***
Discriminant Analysis	***	Fast	Fast	Low	Yes	No

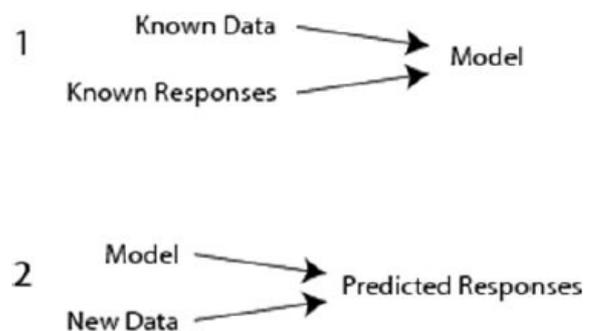


Fig 2: Supervised Learning Algorithm

The two category of supervised learning are:

1. Grouping of responses having only truth values(true or false).
2. Retroversion of responses having real values.

In supervised learning inputs are received but failed to obtain supervised target outputs and rewards from its environment. Though it failed but it is possible for developing a formal framework for unsupervised learning like clustering and dimensionality reduction.

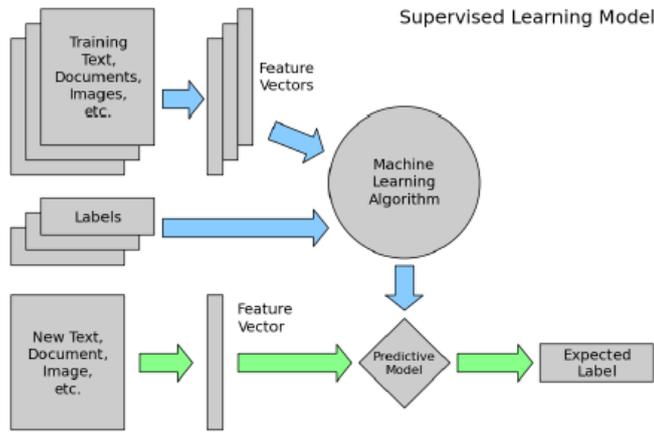


Fig 3: Working Mechanism of Supervised Algorithm

Algorithm for unsupervised learning:

1. HIERARCHIAL CLUSTERING:

Hierarchical clustering is a method of cluster analysis in which we look forward to build a hierarchy of cluster. The aim of this algorithm is to make a multilevel hierarchy of cluster by creating cluster tree.

Inputs: objects represented as vectors

Output: a hierarchy of associations represented as a “Dendogram”

Algorithm:

1. hclust (D,;set of instances):tree
2. var:C;/*set of clusters*/
3. M/*matrix containing distance between 2 clusters*/
4. For each d∈D}do
5. Make a as leaf node in C
6. Done
7. For each pair a,b∈C do
8. Ma,b←d(a,b)
9. Done
10. While(not all instances in one cluster)do
11. Find the most similar pair of cluster in M
12. Merge these two cluster into one cluster
13. Update M to reflect the merge operation
14. Done
15. Return C

2. K-MEANS CLUSTERING

In data mining, a method of vector quantization for cluster analysis is used i.e known as k-means clustering. The aim of

k-means clustering is that the partitioning of n observation belongs to the nearest cluster as prototype.

ALGORITHM:

1. K-means ((X={d1.....dn} Rm,k):2R)
2. C:2R /* μ a set of clusters*/
3. d:RxRm->R/*distance function*/
4. μ:2R->R/*μ computers the mean of a cluster*/
5. select C with k initial centers f1,...,fk
6. while stopping criterion not true do
7. for all clusters cj ∈ C do
8. cj←{di\ϕf1d(di,fj) ≤d(di,f1)}
9. done
10. for all means fj do
11. fj←μ(cj)
12. done

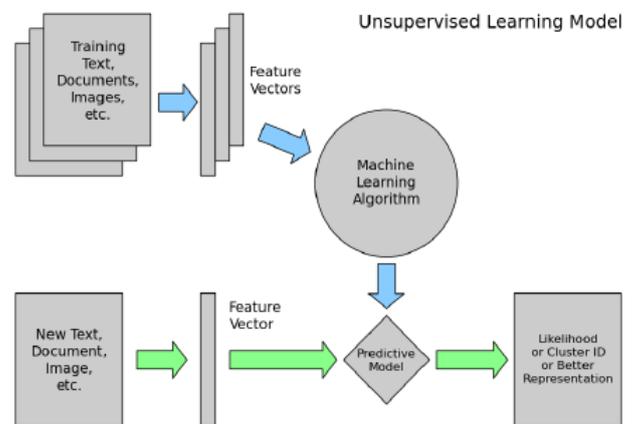


Fig 4: Working Mechanism of Unsupervised Learning Model

CONCLUSION

The investigation of performance measurement of learning algorithm has been studied. This is the complicated query with many aspects. Some issues like analyzing evaluation methods and the metrics which measure performance and a frame work which describe the methods in a structural way. The conclusion that we made by the analysis is that the measurement of classifier performance is calculate by accuracy like in cross validation test. Some general methods are used to evaluate any classifier or any algorithm by the structure of representation, while other methods have restricted to any certain algorithm of representation. The visualization of classifier performance is required because of the method doesn't work like a function returning a performance as result. Measure based evaluation for measuring classifier performance has also been investigated and we provide factual experiments results that strengthen earlier publication of theoretical arguments for measure based evaluation. This experiment was capable of differentiate between classifiers that we were acquainted via accuracy but different in complexity.

REFERENCES

[1] Sally Goldman; Yan Zhou, "Enhancing Supervised Learning with Unlabeled Data", Department of Computer Science, Washington University, St.Louis, MO 63130 USA.

[2] Zoubin Ghahramani, "Unsupervised Learning", Gatsby Computational Neuroscience Unit, University College London, UK.

[3] Rich Caruana; Alexandru Niculescu- Mizil, "An Empirical Comparison of Supervised Learning Algorithms", Department of Computer Science, Cornell University, Ithaca, NY 14853 USA

[4] Niklas Lavesson, "Evaluation and Analysis of Supervised Learning Algorithms and Classifiers", Blekinge Institute of Technology Licentiate Dissertation Series No 2006:04, ISSN 1650-2140, ISBN 91-7295-083-8

[5] Yugowati Praharsi; Shaou-Gang Miaou; Hui-Ming Wee, "Supervised learning approaches and feature selection - a case study in diabetes", International Journal of Data Analysis Techniques and Strategies 2013 - Vol. 5, No.3 pp. 323 - 337

[6] Andrew Ng, "Deep Learning And Unsupervised", "Genetic Learning Algorithms", "Reinforcement Learning and Control", Department of Computer Science, Stanford University, 450 Serra Mall, CA 94305, USA.

[7] Bing Liu, "Supervised Learning", Department of Computer Science, University of Illinois at Chicago (UIC), 851 S. Morgan Street, Chicago

[8] Niklas Lavesson, "Evaluation and Analysis of Supervised Learning Algorithms and Classifiers", Blekinge Institute of Technology Licentiate Dissertation Series No 2006:04, ISSN 1650-2140, ISBN 91-7295-083-8.

[9] Rich Caruana; Alexandru Niculescu- Mizil, "An Empirical Comparison of Supervised Learning Algorithms", Department of Computer Science, Cornell University, Ithaca, NY 14853 USA

[10] Peter Norvig; Stuart Russell, "Artificial Intelligence: A Modern Approach".

BIOGRAPHIES



Mr. A.K. Ratha is working as an assistant professor in department of CSE, Vikash Institute of Technology, Bargarh. He has done MCA in 2007 from Sambalpur University and M.Tech. in CSE with Information Security specialization in 2012 from KIIT University,

Bhubaneswar. He has 10 years of experience in teaching engineering. His research area is Natural Language Processing, Application Programming & Machine Learning.



Ms. Nisha Agrawal is pursuing B.tech(CSE) from Vikash Institute of Technology, Bargarh and presently is in Fourth Year. Her research interests are Database management system and Operating System.



Ms. Amisha Ananya Sikandar is pursuing B.tech(CSE) from Vikash Institute of Technology, Bargarh and presently is in Fourth Year. Her research interests are Operating System and Data Structure