
Machine Learning

Abstract— The problem of learning and decision making is at the core level of argument in biological as well as artificial aspects. So scientist introduced Machine Learning as widely used concept in Artificial Intelligence. It is the concept which teaches machines to detect different patterns and to adapt to new circumstances. Machine Learning can be both experience and explanation based learning. In the field of robotics machine learning plays a vital role, it helps in taking an optimized decision for the machine which eventually increases the efficiency of the machine and more organized way of performing a particular task. Now-a-days the concept of machine learning is used in many applications and is a core concept for intelligent systems which leads to the introduction innovative technology and more advance concepts of artificial thinking. **Keywords**— machine learning, supervised learning, unsupervised learning, algorithms

INTRODUCTION

Learning is considered as a parameter for intelligent machines. Deep understanding would help in taking decisions in a more optimized form and also help then to work in most efficient method. As seeing is intelligence, so learning is also becoming a key to the study of biological and artificial vision. Instead of building heavy machines with explicit programming now different algorithms are being introduced which will help the machine to understand the virtual environment and based on their understanding the machine will take particular decision. This will eventually decrease the number of programming concepts and also machine will become independent and take decisions on their own. Different algorithms are introduced for different types of machines and the decisions taken by them. Designing the algorithm and using it in most appropriate way is the real challenge for the developers and scientists. Pattern recognizing is also a concept in machine learning. Most algorithms use the concept of pattern recognition to make optimized decisions.

As a consequence of this new interest in learning we are experiencing a new era in statistical and functional approximation techniques and their applications to domain such as computer visions. This research paper emphasizes on different types of machine learning algorithms and their most efficient use to make decisions more efficient and complete the task in more optimized form. Different algorithm gives machine different learning experience and adapting other things from the environment. Based on these algorithms the machine takes the decision and performs the specialized tasks. So it is very important for the algorithms to be optimized and complexity should be reduced because more the efficient algorithm more efficient decisions will the machine makes.

Need help with the assignment?

Our professionals are ready to assist with any writing!

GET HELP

Machine Learning algorithms do not totally dependent on nature's bounty for both inspiration and mechanisms. Fundamentally and scientifically these algorithms depends on the data structures used as well as theories of learning cognitive and genetic structures. But still natural procedure for learning gives great exposures for understanding and good scope for Anish Talwar 1IJCS Volume 2 Issue 12, Dec. 2013, Page No.3400-3404 Page 3401 variety of different types of circumstances. Many machine learning algorithm are generally being borrowed from current thinking in cognitive science and neural networks. Overall we can say that learning is defined in terms of improving performance based on some measure. To know whether an agent has learned, we must define a measure of success. The measure is usually not how well the agent performs on the training experiences, but how well the agent performs for new experiences. In this research paper we will consider the two main types of algorithms i.e. supervised & unsupervised learning. . Fig. 1 Diagram representing Machine Learning Mechanism

RELATED WORKS

Sally Goldman et.al [1] proposed the practical learning scenarios where we have small amount of labeled data along with a large pool of unlabeled data and presented a "cotraining" strategy for using the unlabeled data to improve the standard supervised learning algorithms. She assumed that there are two different supervised learning algorithms which both output a hypothesis that defines a partition of instance space for e.g. a decision tree partitions the instance space with one equivalent class defined per tree. She finally concluded that two supervised learning algorithms can be used successfully label data for each other. Zoubin Ghahramani et.al[2] gave a brief overview of unsupervised learning from the perspective of statistical modelling.

According to him unsupervised learning can be motivated from information theoretic and Bayesian principles. He also reviewed the models in unsupervised learning. He further concluded that statistics provides a coherent framework for learning from data and for reasoning under uncertainty and also he mentioned the types of models like Graphical model which played an important role in learning systems for variety of different kinds of data. Rich Caruana et.al [3] has studied various supervised learning methods which were introduced in last decade and provide a large-scale empirical comparison between ten supervised learning methods. These methods include: SVMs, neural nets, logistic regression, naive bayes, memory-based learning, random forests, decision trees, bagged trees, boosted trees and boosted stumps. They also studied and examine the effect that calibrating the models through Platt Scaling and Isotonic Regression has on their performance. They had used various performance based criteria to evaluate the learning methods.

Niklas lavesson et.al [4] answered the fundamental question that how to evaluate and analyse

Need help with the assignment?

Our professionals are ready to assist with any writing!

GET HELP

supervised learning algorithms and classifiers. One conclusion of the analysis is that performance is often only measured in terms of accuracy, e.g., through cross-validation tests. However, some researchers have questioned the validity of using accuracy as the only performance metric. They have given a different approach for evaluation of supervised learning, i.e. Measure functions, a limitation of current measure functions is that they can only handle two-dimensional instance spaces. They present the design and implementation of a generalized multi-dimensional measure function and demonstrate its use through a set of experiments.

The results 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. Finally, they investigate the impact of learning algorithm parameter tuning. Yugowati Praharsi et.al[5] had taken three supervised learning methods such as k-nearest neighbour (k-NN), support vector data description (SVDD) and support vector machine (SVM), as they do not suffer from the problem of introducing a new class, and used them for Data description and Classification. The results show that feature selection based on mean information gain and a standard deviation threshold can be considered as a substitute for forward selection. This indicates that data variation using information gain is an important factor that must be considered in selecting feature subset. Finally, among eight candidate features, glucose level is the most prominent feature for diabetes detection in all classifiers and feature selection methods under consideration. Relevancy measurement in information gain can sort out the most important feature to the least significant one. It can be very useful in medical applications such as defining feature prioritisation for symptom recognition. In this way the analyze the accuracy and working of all the three methods.

PROBLEMS FACED IN LEARNING

Learning is a complex process as lot of decisions are made and also it depends from machine to machine and from algorithm to algorithm, how to understands a particular problem and on understanding the problem how it responds to it. Some of the issues make a complex situation for the machine to respond and react. These problems not only make problem complex it also affects the learning process of the machine. As the machine is dependent on what it perceives, the perception module of the machine should also focus on different types of challenges and environment which it will face, as different input can produce different outputs and the most appropriate and optimize output should be considered by the machine.

Some of the common Anish Talwar 1IJECS Volume 2 Issue 12, Dec. 2013, Page No.3400-3404 Page 3402 problems faced during the learning process are as follows:- Bias- the tendency to prefer one hypothesis over another is called a bias. Consider the agents N and P. Saying that a hypothesis is better than N's or P's hypothesis is not something that is obtained from the data - both N and P accurately predicts all of the data given - but is something external

Need help with the assignment?

Our professionals are ready to assist with any writing!

GET HELP

to the data. Without a bias, an agent will not be able to make any predictions on unseen examples. The hypotheses adopted by P and N disagree on all further examples, and, if a learning agent cannot choose some hypotheses as better, the agent will not be able to resolve this disagreement. To have any inductive process make predictions on unseen data, an agent requires a bias.

What constitutes a good bias is an empirical question about which biases work best in practice; we do not imagine that either P's or N's biases work well in practice. Noise-In most real-world situations, the data are not perfect. Noise exists in the data (some of the features have been assigned the wrong value), there are inadequate features (the features given do not predict the classification), and often there are examples with missing features. One of the important properties of a learning algorithm is its ability to handle noisy data in all of its forms.

Pattern Recognition-This is another type of problem faced in machine learning process. Pattern recognition algorithms generally aim to provide a reasonable answer for all possible inputs and to perform "closest to" matching of the inputs, taking into account their statistical variations. This is different from pattern matching algorithms which match the exact values and dimensions. As algorithms have well-defined values like for mathematical models and shapes like different values for rectangle, square, circle etc. It becomes different for machine to process those inputs which have different values e.g. Consider a ball the shape and pattern can be recognized by the machine, but now when we keep an inflated ball then the pattern would be entirely different and the machine will face problem in recognizing the pattern and the entire process comes to halt. This is the major problem faced by most of the machine learning process and algorithms..

SUPERVISED LEARNING

Supervised learning is an algorithm in which both the inputs and outputs can be perceived. Based on this training data, the algorithm has to generalize such that it is able to correctly respond to all possible inputs. This algorithm is expected to produce correct output for inputs that weren't encountered during training. In supervised learning what has to be learned is specified for each example. Supervised classification occurs when a trainer provides the classification for each example. Supervised learning of actions occurs when the agent is given immediate feedback about the value of each action.

In order to solve a give problem using supervised learning algorithm one has to follow some certain steps:-

1. 1) Determine the type of training examples.
2. 2) Gather a training set.

Need help with the assignment?

Our professionals are ready to assist with any writing!

GET HELP

3. 3) Determine the input feature representation of learned function.
4. 4) Determine the structure of learning function & corresponding learning algorithm.
5. 5) Complete the design and run the learning algorithm on the gather set of data.
6. 6) Evaluate the accuracy of the learned function also the performance of the learning function should be measured and then the performance should be again measured on the set which is different from the training set.

Algorithm Predictive Accuracy Fitting Speed Prediction Speed Memory Usage Easy to Interpret Handle s Categorical Predictors Trees Low Fast Fast Low Yes Yes SVM High Medium * * * No Naïve Bayes Low ** * * * Yes Yes Nearest Neighbor *** Fast*** Medium High No Yes*** Discriminant Analysis **** Fast Fast Low Yes No Anish Talwar 1IJECS Volume 2 Issue 12, Dec. 2013, Page No.3400-3404 Page 3403 Fig. 2 Diagram representing Supervised Learning Algorithm Supervised Learning can be split into two broad categories: 1. Classification of responses that can have just a few values, such as 'true' or 'false'. Classification algorithm applies to nominal, not ordinal response values. 2. Regression for responses that are a real number, such as miles per gallon of a particular car. Characteristics of Algorithm This table shows typical characteristics of the various supervised learning algorithms. The characteristics in any particular case can vary from the listed ones.

Use the table as a guide for your initial choice of algorithms, but be aware that the table can be inaccurate for some problems. Fig. 3 Table showing the characteristics of Supervised Learning Algorithms Discriminant Analysis **** Fast Fast Low Yes No * — SVM prediction speed and memory usage are good if there are few support vectors, but can be poor if there are many support vectors. When you use a kernel function, it can be difficult to interpret how SVM classifies data, though the default linear scheme is easy to interpret. ** — Naive Bayes speed and memory usage are good for simple distributions, but can be poor for kernel distributions and large data sets. *** — Nearest Neighbor usually has good predictions in low dimensions, but can have poor predictions in high dimensions. For linear search, Nearest Neighbor does not perform any fitting. For kd-trees, Nearest Neighbor does perform fitting. Nearest Neighbor can have either continuous or categorical predictors, but not both. **** — Discriminant Analysis is accurate when the modeling assumptions are satisfied (multivariate normal by class). Otherwise, the predictive accuracy varies. Fig. 4 This block-diagram shows the working mechanism of Supervised Learning.

SUPERVISED LEARNING

In unsupervised learning the machine simply receives the input x_1, x_2, \dots but obtains neither supervised target outputs, nor rewards from its environment. But it is possible to develop a formal framework for unsupervised learning based on the notion that the machine's goal is to build representations of the input that can be used for decision making, predicting future inputs,

Need help with the assignment?

Our professionals are ready to assist with any writing!

[GET HELP](#)

efficiently communicating the inputs to another machine, etc. Example of unsupervised learning is clustering and dimensionality reduction. Some algorithms for unsupervised learning are as follows: 1. Hierarchical clustering This algorithm builds a multilevel hierarchy of cluster by creating a cluster tree. Inputs: objects represented as vectors Outputs: a hierarchy of associations represented as a “dendrogram”. Algorithm: 1. $hclust(D: \text{set of instances}): \text{tree}$ 2. var: $C, /* \text{set of clusters} */$ 3. $M /* \text{matrix containing distances between pairs of cluster} */$ 4. for each $d \in D$ do 5.

Make d a leaf node in C 6. done 7. for each pair $a, b \in C$ do 8. $M_{a,b} = d(a, b)$ 9. done 10. while(not all instances in one cluster) do 11. Find the most similar pair of clusters in M 12. Merge these two clusters into one cluster. 13. Update M to reflect the merge operation. 14. done 15. return C K-means structuring Anish Talwar 1IJECS Volume 2 Issue 12, Dec. 2013, Page No.3400-3404 Page 3404 In this algorithm we have to first select the number of cluster in advance, they might converge to a local minimum. K-means can be seen as a specialization of the expectation maximization (EM) algorithm. It is more efficient (lower computational complexity) than hierarchical clustering. Algorithm: 1. K-means ($(X = \{d_1, \dots, d_n\} \in R^m, k): 2R$) 2. $C: 2R /* \mu \text{ a set of clusters} */$ 3. $d = R^m \times R^m \rightarrow R /* \text{distance function} */$ 4. $\mu: 2R \rightarrow R /* \mu \text{ computes the mean of a cluster} */$ 5. select C with k initial centers f_1, \dots, f_k 6. while stopping criterion not true do 7. for all clusters $c_j \in C$ do 8. $c_j = \{d_i \mid d(d_i, f_j) < d(d_i, f_l)\}$ 9. done 10. for all means f_j do 11. f_j

Need help with the assignment?

Our professionals are ready to assist with any writing!

[GET HELP](#)