ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

A Review on Different Classification Techniques

A. Tamilmozhi [1], Dr.N.Sujatha [2]

Research Scholar, PG & Research Department of Computer Science, Raja Doraisingam Govt. Arts College, Sivagangai, TamilNadu, India [1]

Assistant Professor, PG & Research Department of Computer Science, Raja Doraisingam Govt. Arts College, Sivagangai, TamilNadu, India [2]

ABSTRACT: In this paper distinctive classification methods of Data mining are looked at utilizing various datasets from UCI machine learning . Precision and time required for execution by every method is noticed. The Data mining alludes to learning from large volume of information. Classification is an essential data mining method with wide applications. It arranges information of different sorts. Classification is utilized as a part of each field of our life. Classification is utilized to order everything in a group of information into one of predefined set of classes or bunches. This work has been done on Labor dataset by applying some of the classification algorithms such as: J48, Multilayer Perceptron, and Bayes Net algorithms. Experimental results are given to evaluate the performance of each algorithm.

KEYWORDS: Data mining, classification, J48, Multilayer Perceptron, and Bayes Net.

I. INTRODUCTION

Data mining involves the use of various sophisticated data analysis tools for discovering previously unknown, valid patterns and relationships in huge data set. These tools are nothing but the machine learning methods, statistical models and mathematical algorithm. Data mining consists of more than collection and managing the data, it also includes analysis and prediction. Classification technique in data mining is capable of processing a wider variety of data than regression and is growing in popularity.

There are number of applications for Machine Learning (ML), the most significant of which is data mining. The term Data Mining, also known as Knowledge Discovery in Databases (KDD) refers to the nontrivial extraction of implicit, potentially useful and previously unknown information from data in databases [1]. There are several data mining techniques are association, classification, pattern recognition and clustering, [2]. Classification and association are the popular techniques used to predict user interest and relationship between those data items which has been used by users [3, 4].

In this paper, labor dataset is taken which is taken from the UCI machine learning repository. It contains labor details of a company in the name of 17 attributes such as: duration, wage increase-first-year, wage increase-second-year, wage increase-third-year, cost of living adjustments,working hours,pension,standby pay,shift differential,education allowance,statutory holidays,vacation,long-term disability assistance,contribution to dental plan,bereavement assistance,contribution to health plan and class. The classification algorithms like: J48 Decision tree, Bayesian network, and Multilayer Perceptron are applied on this labor dataset in order to evaluate its performance.

II. CLASSIFICATION TECHNIQUES

2.1. J48

It can be called as advanced execution of the C4.5 or enhanced form of the C4.5.[5]. The yield given by J48 is the Choice tree. A Choice tree is same as that of the tree structure having diverse nodes, for example, root node, moderate nodes and leaf node. [6]. Every node in the tree contains a choice and that choice prompts our outcome as name is choice tree. Choice tree separate the information space of information set into fundamentally unrelated territories, where every range having a mark, a worth or an activity to portray or expand its information focuses. Part paradigm is

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

utilized as a part of choice tree to figure which credit is the best to part that partition tree of the preparation information that achieves a specific node. [7].

A. Algorithm

1.ID3(D, Attributes,Target)
t=createNode()
2. label(t) = mostCommonClass(D, Target)
3. IF ∀hx, c(x)i ∈ D : c(x) = c THEN return(t) ENDIF
4. IF Attributes = ∅ THEN return(t) ENDIF
5. A∗ = argmaxA∈Attributes(informationGain(D,
6. FOREACH a ∈ A∗ DO Da = {(x, c(x)) ∈ D : x|A∗ = a} IF Da = ∅ THEN t 0 = createNode() label(t 0) = mostCommonClass(D, Target) createEdge(t, a, t0) ELSE createEdge(t, a, ID3(Da,Attributes \ {A∗}, Target)) ENDIF ENDDO
7. return(t)

B. **Flow Chart**

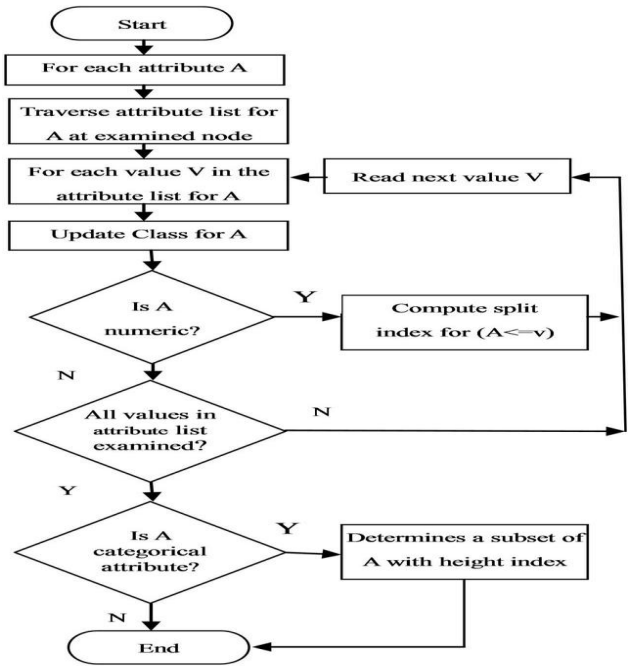

Figure 1: Flow chart of J48

C. Advantages

1. Decision trees are easy to understand;
2. Decision trees are easily converted to a set of production rules;
3. Decision trees can classify both categorical and numerical data, but the output attribute must be categorical;
4. There are no a priori assumptions about the nature of the data.

D. Disadvantages

1. Decision boundaries are rectilinear.
2. Small variations in the data can imply that very different looking trees are generated.
3. A sub-tree can be replicated several times.
4. Error-prone with too many classes.
5. Not good for predicting the value of a continuous class attribute.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

2.2. Multi Layer Perceptron

Multi Layer Perceptron **c**an be characterized as Neural Network and Artificial knowledge without capability. A Multi Layer perceptron (MLP) is a feedforward neural system with one or more layers in the middle of info and yield layer.[8]. Essentially there are three layers: information layer, concealed layer and yield layer. Concealed layer might be more than one. Every neuron (node) in every layer is associated with each neuron (node) in the neighboring layers. The preparation or testing vectors are joined with the info layer, and facilitate handled by the covered up and yield layers.

Back Propagation Algorithm Back propagation is a form of supervised learning for multi-layer nets, also known as the generalized delta rule. Error data at the output layer is back propagated to earlier ones, allowing incoming weights to these layers to be updated. It is most often used as training algorithm in current neural network applications. The back propagation algorithm is an involved mathematical tool; however, execution of the training equations is based on iterative processes, and thus is easily implementable on a computer.[9].

During the training session of the network, a pair of patterns is called input pattern and desired pattern. At the output layer, the difference between the actual and target outputs yields an error signal. This error signal depends on the values of the weights of the neurons in each layer. This error is minimized, and during this process new values for the weights are obtained. The speed and accuracy of the learning process that is, the process of updating the weights-also depends on a factor, known as the learning rate.

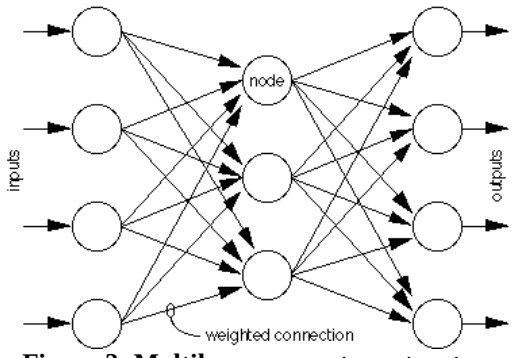

Figure 2: Multilayer perceptron structure

A. Algorithm:

1: choose an initial weight vector $\vec{w}$

2: initialize minimization approach

3: while error did not converge do

4: for all $(\vec{x}, \vec{d}) \in D$ do

5: apply $\vec{x}$ to network and calculate the network output

6: calculate $\frac{\partial e(\vec{x})}{\partial w_{i,j}}$ for all weights

7: end for

8: calculate $\frac{\partial E(D)}{\partial w_{i,j}}$ for all weights summing over all training patterns

9: perform one update step of the minimization approach

10: end while

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

B. Flow Chart:

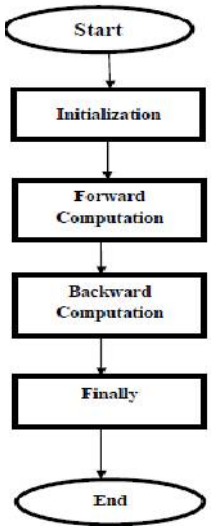

Figure 3: Flow of Multi Layer Perceptron

C. Advantages

1. Fault Tolerance Easy To Identify
2. Speech synthesis in Main Approach
3. Financial applications
4. Pattern Recognition

D. Disadvantages

1. Relatively insensitive to outliers (wild points).
2. Slow Process is Main Drawback.
3. Low memory space to store the model.

2.3 BayesNet classifier

BayesNet classifier depends on the bayes theorm. Along these lines, in BayesNet classifier restrictive likelihood on every node is ascertained first and after that a Bayesian System gets shaped. Bayesian System is nothing be that as it may, a coordinated non-cyclic chart. The supposition made in BayesNet is, that all qualities are ostensible and there are no missing values any such esteem supplanted all inclusive. Hereditary Calculation and K2 such an alternate sorts of calculations are utilized to assess restrictive likelihood in BayesNet. In BayesNet, the yield of can be imagined as far as chart. [10].

A. Algorithm:

1.Input: A set S of labeled examples, a set X of variables.
Output: A full BN B. 1. B = empty.
2. Partition the training data S into |C| subsets Sc by the class value c.
3. For each training data set Sc
Compute the mutual information M(Xi ; Xj) and the structure penalty ϕ(Xi , Xj) between each pair of variables Xi and Xj based on Equation 1 and
Equation 2
Compute W(Xi) for each variable Xi based on Equation 3.
For all variables Xi in X – Add all the variables Xj with W(Xj) > W(Xi) to the parent set ΠXi ofXi . – Add arcs from all the variables Xj in ΠXi to Xi .
add the resulting network Bc to B.
4.Return B.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

B. Flow Chart:

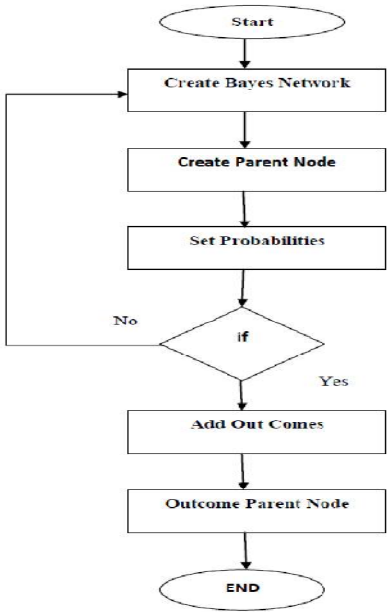

Figure 4: Flow of BayesNet classifier

C. Advantages

1. Bayesian learning methods interpolate all the way to pure engineering.
2. This is significantly helpful when working on the "think harder" part of a solution.
3. Bayesian learning involves specifying a prior and integration, two activities which seem to be universally useful. (See intuitions).

D. Disadvantages

1. It turns out that specifying a prior is extremely difficult.
2. Let's suppose we could accurately specify a prior over every air molecule in a room. Even then, computing a posterior may be extremely difficult. This difficulty implies that computational approximation is required.
3. The "think harder" part of the Bayesian research program is (in some sense) a "Bayesian employment" act.

III. RESULTS AND DISCUSSION

A comparison of classifiers for different datasets based on the accuracy and time taken for execution is made. Accuracy is defined as the no of instances classified correctly. It is observed from table 4.1 that, Multilayer perceptron performed well with Labor dataset.

Algorithm	Correctly Classified (%)	In correctly Classified (%)	Time Execution (sec)
J48	73.68	26.13	0.02
Multilayer Perceptron	92.98	7.01	0.05
Bayes Net	87.71	12.28	0.02

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

Table: 4.1 Comparison of correctly and incorrectly classified for various Classification algorithms in Labor Dataset

Algorithm	Precision	Recall	F-Measure
J48	0.748	0.737	0.74
Multilayer perceptron	0.934	0.93	0.931
Bayes net	0.886	0.877	0.879

Table: 4.2 performance measures for various classification algorithms in Labor Dataset

IV. CONCLUSION

Thus we evaluate the performance in terms of classification accuracy of J48 and Multilayer Perceptron and Bayes Net algorithms using various accuracy and time measures. Accuracy has been measured on each classification algorithm which is tabulated from here. It is clearly shows that Multilayer Perceptron provides better performance than J48 and Bayes Net algorithms in labour dataset. Our future work will concentrate on change of classification Method in this way enhancing the proficiency of order in a diminished time. Additionally a mix of characterization systems will be used to enhance the performance.

REFERENCES

[1].Usama Fayyad, Gregory Piatetsky-Shapiro, Padhraic Smyth, and Ramasamy Uthurasamy, "Advances in Knowledge Discovery and Data Mining", AAAI Press/ The MIT Press, 1996.
[2]. Michael Berry and Gordon Linoff, "Data Mining Techniques (For Marketing, Sales, and Customer Support), John Wiley & Sons, 1997
[3]. K. Rajeswari, Dr. V. Vaithiyannathan,(2012) “Mining Association Rules Using Hash Table”, International Journal of Computer Applications(9132-3320).
[4]. Rajeswari, Dr. V. Vaithiyannathan,(2011), “ Heart Disease Diagnosis: An Efficient Decision Support System Based on Fuzzy Logic and Genetic Algorithm”, International Journal of Decision Sciences, Risk and Management by Inderscience Publications. ISSN: 1753-7169.
[5]. J. Ross Quinlan, "C4.5: Programs for Machine Learning", Morgan Kaufmann Publishers, 1993
[6].Ozer, Patrick. "Data Mining Algorithms for Classification."(January 2008): 27.Web. 5 May 2013.[16].
[7].Gholap, jay"Performance tuning of j48 algorithm for prediction of soil fertility."5.web. 2 May 2013
[8]. Fine, T. L. 1999. Feedforward Neural NetworkMethodology, 3rd ed. New York: Springer-Verlag.
[9]. *Rumelhart, David E.; Hinton, Geoffrey E.; Williams, Ronald J. (8 October 1986). "Learning representations by back-propagating errors".*
[10].Friedman, N., D Geiger, M Goldszmidt “Bayesian Networks Classifiers” *Machine Learning*, 29, 131–163 (1997).