A Survey of Classification Methods Utilizing Decision Trees

Prerna Kapoor[#1], Reena Rani[*2]

[#]M.TECH&Computer Science & Engineering Department&
JMIT, Radaur/ Kurukshetra University, India
[*]Assistant Professor & Computer Science & Engineering Department&
JMIT, Radaur/ Kurukshetra University, India

***Abstract*—Therecognition of outlines and the invention of decision rules from data is one of the challenging setbacks in discovering and learning.When continuous attributes are involved in the process the attributes should be discretized with threshold values or with various other standardizing methods. Decision tree induction algorithms craft decision trees by recursively partitioning the input space. Hence, a rule tree is obtained by traversal from the origin node to every single leaf node in the tree. The decision trees can be fiercely embodied as a set of decision laws (if-then-else rules) to assist the understanding. Inductive discovering methods craft such decision trees, frequently established on heuristicdata or statistical probability concerning attributes. This paper is about Decision Trees algorithms and their implementation mainly C4.5.**

***Keywords*—Machine Learning, Data mining, Decision trees, C4.5, J48.**

I. Introduction

Decision trees are a simple, but powerful form of multiple variable analysis. Theyprovide unique capabilities to supplement, complement, and substitute for

- traditional statistical forms of analysis
- a variety of data mining tools and techniques
- recently developed multidimensional forms of reporting and analysis found in the field of business intelligence

Decision trees are produced by algorithms that recognize assorted methods of dividing a data set into branch-like segments. These segments form an inverted decision tree that originates alongside the origin node at the top of the tree. The object of analysis is imitated in this origin node as an easy, one-dimensional display in the decision tree interface. The term of the field of data that is the object of analysis is normally displayed, alongside the range or allocation of the benefits that are encompassed in that field. The display of this node reflects all the data set records, fields, and fieldvalues that are discovered in the object of analysis. The invention of the decision law to form the divisions or segments underneath the origin node is established on a method that extracts the connection amid the object of analysis and one or extra fields that assist as input fields to craft the divisions or segments. The benefits in the input field are utilized to guesstimate the probable worth in the target field. The target field is additionally known as a consequence, reply, or reliantfield or variable.

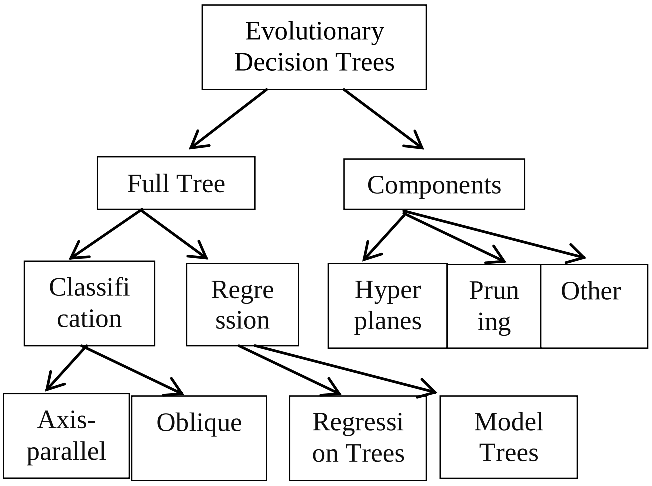

Fig. 1Taxonomy of evolutionary algorithms for decision tree induction.

II. Classification

There are various classifications which are related to Decision Trees [1]:

A. Axis-Parallel Decision Trees

Axis-parallel decision trees are the most public kind discovered in the works, generally because this kind of tree is normally far easier to clarify than an oblique tree. We splitour analysis on axis-parallel decision trees according to the main steps of the evolutionary process. That is, we examine how resolutions are encoded; that methods are utilized for initializing the population of decision trees; the most public strategies for fitness evaluation; the genetic operators that are projected to evolve individuals; and supplementary connected issues.

1) Solution Encoding:

Some terminology subjects that are normally dictated according to the EA (evolutionary algorithm) resolution encoding scheme. Nomenclature aside, decision tree encoding is normally whichever tree-based or non-tree based. We comment on both next. Tree-based encoding is the most public way for coding people in EAs for decision tree induction, and it seems a usual choice after we are dealing alongside decision trees. The competitive co-evolution for decision tree induction and uses a tree-encoding scheme. The arrangement sketches binary decision trees whereas every single node is embodied by a 4-tuple.Each constituent is a numeric worth that can be adjusted across the evolutionary process.

2) *Population Initialization:*

An EA's early populace has to furnish plenty diversity of people so that the genetic operators can find for resolutions in an extra comprehensive search-space, circumventing innate optima. Nonetheless, a colossal search-space could consequence in extremely sluggish convergence, stopping the EA from discovering a near-optimal solution. In this case, task reliant vision constraints could speed-up convergence by circumventing the find in "dead zones" of the resolution space. It is clear that there is a slender line amid the precise number of diversification for circumventing innate optima and task-dependent vision constraints for speeding-up convergence.

3) *Fitness Evaluation Methods:*

Evolutionary decision tree induction algorithms can be roughly split into two threads considering fitness evaluation: single-objective optimization and multi-objective optimization. EAs that present single-objective optimization use a solitary compute to escort the find for near-optimal solutions. The most public compute for assessing entities in evolutionary algorithms for decision tree induction is association accuracy:

$$acc = \frac{c}{m}$$

where c is the number of accurately categorized instances and m is the finished number of instances.

4) *Selection Methods and Genetic Operators:*

Selection is the procedure that chooses that people will experience crossover and mutation. In evolutionary induction of decision trees, the most oftentimes utilized way for selection is tournament selection. One more accepted choice in EAs for decision tree induction is the roulette wheel selection. A less-common selection method in EAs for decision tree induction is rank-based selection. Two operators normally utilized to evolve a populace of people are crossover and mutation. In EAs for decision tree induction, crossover is normally given in two disparate methods according to the individual representation. For fixed-length binary thread encoding, it is a public way to present the well-known 1-point crossover.

5) *Parameter Setting:*

The parameter benefits of an EA can mainly impact whether the algorithm will find an adjacent optimum resolution, and whether it will find such a resolution efficiently. The most public parameters in EAs for decision tree induction are populace size, number of generations, probabilities of request of disparate genetic operators and maximum size of decision trees at initialization or across the evolutionary process. In exercise, countless preliminary runs are normally needed in order to tune these parameters. Though, most authors favor to present a set of default parameter valuespursued by a sentence like "parameter benefits were empirically defined".

B. *Oblique Decision Trees*

Oblique decision trees, additionally denoted to as (non-) linear decision trees, are a public alternative to the established axis parallel approach. Oblique decision trees are normally far tinier and frequently extra precise than axis-parallel decision trees, nevertheless at the price of extra computational power and defeat of comprehensibility. In oblique decision trees, hyper plane divides the feature space into two distinct spans.

III. DECISION TREE ALGORITHM

A. *C4.5*

The C4.5 algorithm generates a decision tree for the given data by recursively dividing that data. The decision tree grows employing Depth-first strategy. The C4.5 algorithm considers all the probable examinations that can split the data and selects anexamination that gives the best data gain. This examination removes ID3's bias in favor of expansive decision trees. For every single discrete attribute, one examination is utilized to produce countless consequencesas the number of different benefits of the attribute. For every single constant attribute, the data is sorted, and the entropy gain is computed established on binary cuts on every single different worth in one scan of the sorted data. This procedure is recapped for all constant attributes. The C4.5 algorithm permits pruning of the emerging decision trees. This increases the error rates on the training data, but vitally, cuts the error rates on the unseen assessing data. The C4.5 algorithm can additionally deal alongside numeric qualities, missing benefits, and loud data. It has the pursuing gains and disadvantages:

Advantages:

- C4.5 can handle both continuous and discrete attributes. In order to handle continuous attributes, it creates a threshold and then splits the list into those whose attribute value is above the threshold and those that are less than or equal to it.
- C4.5 allows attribute values to be marked as?For missing. Missing attribute values are simply not used in gain and entropy calculations.
- C4.5 goes back through the tree once it's been created and attempts to remove branches that do not help by replacing them with leaf nodes.

Disadvantages:

- C4.5 constructs empty branches; it is the most crucial step for rule generation in C4.5.We have found many nodes with zero values or close to zero values. These values neither contribute to generate rules nor help to construct any class for classification task. Rather it makes the tree bigger and more complex.
- Over fitting happens when algorithm model picks up data with uncommon characteristics. Generally C4.5 algorithm constructs trees and grows it branches 'just deep enough to perfectly classify the training examples'.

- Susceptible to noise.

1) *Decision Trees and C4.5*

A decision tree is a classifier which conducts recursive partition over the instance space. A typical decision tree is composed of internal nodes, edges and leaf nodes. Each internal node is called decision node representing a test on an attribute or a subset of attributes, and each edge is labeled with a specific value or range of value of the input attributes. In this way, internal nodes associated with their edges split the instance space into two or more partitions. Each leaf node is a terminal node of the tree with a class label. For example, Figure 2 provides an illustration of a basic decision tree, where circle means decision node and square means leaf node. In this example, we have three splitting attributes, i.e., age, gender and criteria 3, along with two class labels, i.e., YES and NO. Each path from the root node to leaf node forms a classification rule.

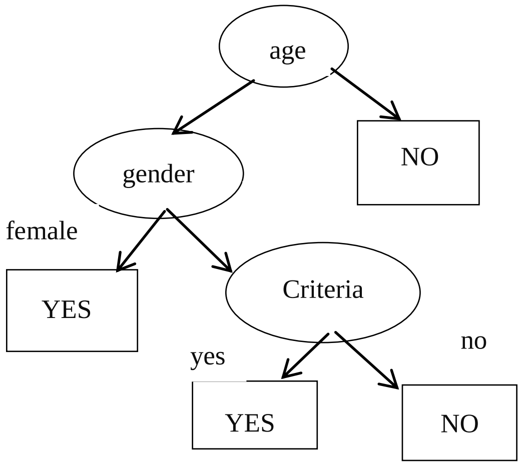

Fig.2 Illustration of Decision Tree

The general process of building a decision tree is as follows. Given a set of training data, apply a measurement function onto all attributes to find a best splitting attribute. Once the splitting attribute is determined, the instance space is partitioned into several parts. Within each partition, if all training instances belong to one single class, the algorithm terminates. Otherwise, the splitting process will be recursively performed until the whole partition is assigned to the same class. Once a decision tree is built, classification rules can be easily generated, which can be used for classification of new instances with unknown class labels.

C4.5 [4] is a standard algorithm for inducing classification rules in the form of decision tree. As an extension of ID3 [5], the default criteria of choosing splitting attributes in C4.5 is information gain ratio. Instead of using information gain as that in ID3, information gain ratio avoids the bias of selecting attributes with many values.

```
Algorithm 1 C4.5(T)
Input: training dataset T; attributes $S$.
Output: decision tree $Tree$.
if $T$ is NULL then
        return failure
end if
if $S$ is NULL then
        return $Tree$ as a single node with mostfrequent class
label in $T$
end if
if $|S| = 1$ then
     return $Tree$ as a single node $S$
        end if
        set $Tree = \{\}$
for $a \in S$ do
set$Info(a,T) = 0$ and $SplitInfo(a,T) = 0$
compute $Entropy\ (a)$
for $v \in values(a,T)$ do
set $T_{a,v}$ as the subset of $T$ with attribute $a = v$
$Info(a,T) += \frac{|T_{a,v}|}{|T_a|} Entropy\ (a_v)$
$SplitInfo(a,T) += -\frac{|T_{a,v}|}{|T_a|} log \frac{|T_{a,v}|}{|T_a|}$
end for
$Gain(a, T) = Entropy(a) - Info(a, T)$
$GainRatio(a,T) = \frac{Gain\ (a,T)}{SplitInfo(a,T)}$
end for
set $a_{best} = argmax\{GainRatio(a,T)\}$
attach $a_{bcst}$ into $Tree$
for $v \in values\ (a_{best}, T)$ do
call C4.5($T_{a,v}$)
end for
return $Tree$
```

Fig.3 C4.5 Algorithm Description

Let C denote the number of classes, and p(S, j) is the proportion of instances in S that are assigned to j-th class. Therefore, the entropy of attribute S is calculated as:

$$\text{Entropy (s)} = -\sum_{j=1}^{c} p(S,j) \times \log p(S,j)$$

Accordingly, the information gain by a training dataset T is defined as:

$$\text{Gain(S,T)} = \text{Entropy(S)} - \sum_{v \in Values(T_S)} \frac{|T_{S,v}|}{|T_S|} Entropy(S_v)$$

where Values (T_s) is the set of values of S in T, T_s is the subset of T induced by S , and $T_{S,v}$, is the subset of T in which attribute S has a value of v.

Therefore, the information gain ratio of attribute S is defined as:

$$\text{GainRatio(S,T)} = \frac{Gain(S,T)}{SplitInfo(S,T)}$$

whereSplitInfo(S,T) is calculated as:

$$\text{SplitInfo(S,T)} = -\Sigma_{v \in Values(T_S)} \frac{|T_{S,v}|}{|T_S|} \times log \frac{|T_{S,v}|}{|T_S|}$$

The whole process of C4.5 algorithm is described in Algorithm 1. The information gain ratio criteria computation is performed in above equations, and a recursive function call is done.

B. J48

J48 is an implementation of C4.5 algorithm [1]. There are two methods in pruning proposed by J48 early are recognized as sub tree substitute, it work by substituting nodes in decision tree alongside leaf. Basically by cut the number of inspection alongside precise path. It works alongside the procedure of commencing from leaves that finished industrialized tree and do a reversing in the direction of the root. The subsequent kind requested in J48 is sub tree rising by advanced nodes upwards in the direction of the origin of tree and additionally substituting supplementary nodes on the alike way. On data assessing this algorithm will emphasized on dividing dataset and by selecting a check that will give best consequence in data gain. In discrete qualities as well, these algorithms ponder a check alongside a consequence of countless as the number of disparate benefits and examination binary attribute for every single attribute will tolerate to produce in disparate benefits every single attribute will be considered.

```
Algorithm J48:

INPUT: D        //Training data
OUTPUT: T       //Decision tree

DTBUILD(*D)
{
T=NULL;
T=Create root node and label with splitting attribute;
T=Add arc to root node for each split predicate and label;
For each arc do
        D= Database created by applying splitting predicate
        to D;
        If stopping point reached for this path, then
                T'= create leaf node and label with
                appropriate class;
        Else
                T'=DTBUILD(D);
        T=add T' to arc;
}
```

C. REDUCED ERROR PRUNING

Basically DecreasedError Pruning Tree ("REPT") is fast decision tree discovering and it builds a decision tree established on the data gain or cutting the variance [2]. The frank of pruning of this algorithm is it utilized REP alongside back above fitting. It kindly sorts benefits for numerical attribute after and it grasping the missing benefits alongside embedded method by C4.5 in fractional instances. In this algorithm we can discern it utilized the method from C4.5 and the frank REP additionally count in it process.

IV. RELATED WORK

Rodrigo C. Barros et al., 2012 [1] This paper presents a survey of evolutionary algorithms projected for decision tree induction. In this context, most of the paper focuses on ways that evolve decision trees as an alternate heuristics to the established top-down divide-and-conquer approach. Additionally, they present a little alternative method that makes use of evolutionary algorithms to enhance particular constituents of decision tree classifiers. The paper early contributions are the following. First, it provides an uptodate overview that is fully concentrated on evolutionary algorithms and decision trees and does not ponder on each specific evolutionary approach. Second, it provides a taxonomy that addresses works that evolve decision trees and works that design decision tree constituents employing evolutionary algorithms.

W. NorHaizan W. Mohamed et al., 2012 [2] In this paper Decision tree is one of the most accepted and effectual method in data mining. This method has been instituted and well-explored by countless researchers. Though, a little decision tree algorithms could produce a colossal construction of tree size and it is tough to understand. Furthermore, misclassification of data frequently occurs in discovering process. Therefore, a decision tree algorithm that can produce an easy tree construction alongside elevated accuracy in word of association rate is a demand to work alongside huge volume of data. Pruning methods have been given to cut the intricacy of tree construction lacking cut the accuracy of classification. One of pruning methods is the Reduced Error Pruning (REP).

Tina R. Patil et al., 2013 [3] In this paper Association is an vital data excavating method alongside colossal requests to categorize the assorted kinds of data utilized in nearly every singlefield of our life. Association is utilized to categorize the item according to the features of the item alongside respect to the predefined set of classes. This paper sitea light on presentation evaluation established on the correct and incorrect instances of data association employing Naïve Bayes and J48 association algorithm. Naive Bayes algorithm is established on probability and j48 algorithm is established on decision tree.

Smith Tsang et al., 2011 [4] Instituted decision tree classifiers work alongside data whose benefits are recognized and precise. They spread such classifiers to grasp data alongside tentative information. Worth uncertainty arises in countless requests across the data collection process. Example origins of uncertainty contain measurement/quantization errors, data staleness, and several recapped measurements. With uncertainty, the worth of a data item is frequently embodied not by one solitary worth, but by several benefits growing a probability distribution. Rather than abstracting tentative data by statistical derivatives (such as mean and median), they notice that the accuracy of a decision tree classifier can be far enhanced if the “complete information” of a data item (taking into report the probability density function (pdf)) is utilized. They spread classical decision tree constructing algorithms to grasp data tuples alongside tentative values. Comprehensive examinations have been led that display that the emerging classifiers are extra precise than those employing worth averages.

Rodrigo Coelho Barros et al., 2011 [5] Decision tree induction is one of the most employed methods to extract knowledge from data, since the representation of knowledge is very intuitive and easily understandable by humans. The most successful strategy for inducing decision trees, the greedy top-down approach, has been continuously improved by researchers over the years. This work, following recent breakthroughs in the automatic design of machine learning algorithms, proposes two different approaches for automatically generating generic decision tree induction algorithms. Both approaches are based on the evolutionary algorithms paradigm, which improves solutions based on metaphors of biological processes. They also propose guidelines to design interesting fitness functions for these evolutionary algorithms, which take into account the requirements and needs of the end-user.

Raj Kumar et al., 2012 [6] In this paper association is an ideal discovering procedure that is utilized for portioning the data into disparate classes according to a little constrains. In supplementary words they can say that association is procedure of generalizing the data according to disparate instances. Countless main kinds of association algorithms encompassing C4.5, k-nearest acquaintance classifier, Naive Bayes, SVM, Apriori, and Ada Boost. These papers furnish an inclusive survey of disparate association algorithms.

A.S. Galathiya et al., 2012 [7] In this analysis work, Analogy is made amid ID3, C4.5 and C5.0. Amid these classifiers C5.0 gives extra precise and effectual output alongside moderately elevated speed. Recollection custom to store the law set in case of the C5.0 classifier is less as it generates tinier decision tree. This analysis work supports elevated accuracy, good speed and low recollection custom as counseled arrangement is employing C5.0 as the center classifier. The association procedure here has low recollection custom difference to supplementary methods because it generates less rules. Accuracy is elevated as error rate is low on unseen cases. And it is fast due to producing pruned trees.

Susan Lomax et al., 2013 [8] In this paper the past decade has perceived a momentous attention on the setback of instigating decision trees that take hold report of prices of misclassification and prices of buying the features utilized for decision making. This survey identifies above 50 algorithms encompassing ways that are manage adaptations of accuracy established methods, use genetic algorithms, use anytime methods and use boosting and bagging. The survey brings jointly these disparate studies and novel ways to cost-sensitive decision tree discovering, provides a functional taxonomy, a past timeline of how the field has industrialized and ought to furnish a functional reference point for upcoming analysis in this field.

Mohammed Abdul Khaleel et al., 2013 [9] In this paper in the last decade there has been rising custom of data excavating methods on health data for discovering functional trends or outlines that are utilized in diagnosis and decision making. Data excavating methods such as clustering, association, regression, association law excavating, CART (Classification and Regression Tree) are extensively utilized in healthcare domain. Data excavating algorithms, after appropriately utilized, are capable of enhancing the quality of forecast, diagnosis and illness classification. The main focus of this paper is to examine data excavating methods needed for health data excavating exceptionally to notice innately recurrent illnesses such as heart ailments, lung cancer, and breast cancer and so on. They assess the data excavating methods for discovering innately recurrent outlines in words of price, presentation, speed and accuracy. They additionally difference data excavating methods alongside standard methods.

AnujaPriyama et al., 2013 [10] In this paper at the present period, the number of data stored in educational database is rising swiftly. These databases encompass hidden data for enhancement of student’s performance. Association of data objects is a data excavating and vision association method utilized in gathering comparable data objects together. There are countless association algorithms obtainable in works but decision tree is the most usually utilized because of its ease of killing and easier to comprehend contrasted to supplementary association algorithms. The ID3, C4.5 and CART decision tree algorithms proceeding requested on the data of students to forecast their performance. But all these are utilized merely for tiny data set and needed that all or a serving of the whole dataset stay perpetually in memory.

Richa Sharma et al., 2013 [11] In this paper an endeavor has been made to develop a decision tree association (DTC) algorithm for association of remotely detected satellite data (Land sat TM) employing open basis support. The decision tree is crafted by recursively partitioning the spectral allocation of the training dataset employing WEKA, open basis data excavating software. The categorized picture is contrasted alongside the picture categorized employing classical ISODATA clustering and Maximum Likelihood Classifier (MLC) algorithms. Association consequence established on DTC method endowed larger discernible portrayal than outcome produced by ISODATA clustering or by MLC algorithms.

LeszekRutkowski et al., 2013 [12] In this paper in excavating data streams the most accepted instrument is the Hoeffding tree algorithm. It uses the Hoeffding’s attached to ascertain the smallest number of examples demanded at a node to select a dividing attribute. In works the alike Hoeffding's attached was utilized for each evaluation purpose (heuristic measure), e.g. data gain or Gini index. In this paper it is shown that the Hoeffding's inequality is not appropriate to resolve the underlying problem. They clarify two theorems giving the McDiarmid's attached for both the data gain, utilized in ID3 algorithm, and for Gini index, utilized in CART algorithm. The outcome of the paper promise that a decision tree discovering arrangement, requested to data streams and established on the McDiarmid'sattached, has the property that its output is nearly identical to that of a standard learner. The consequencesof the paper have an outstanding encounter on the state of the fine art of excavating data streams and assorted industrialized so distant methods and algorithms ought to be reconsidered.

Nirmal Kumar et al., 2013 [13] In this paper Land skill association (LCC) of a dirt chart constituent is pursued for sustainable use, association and conservation practices. Elevated speed,elevated precision and easy producing of laws by contraption discovering algorithms can be utilized to craft pre-defined laws for LCC of dirt chart constituents in growing decision prop arrangements for field use arranging of an area. Decision tree (DT) is one of the most accepted association algorithms presently in contraption discovering and data mining. Creation of Best Early Tree (BF Tree) from qualitative dirt survey data for LCC described in reconnaissance dirt survey data of Wardha district, Maharashtra has been clarified in the present discover alongside dirt depth, hill, and erosion as qualities for LCC. A 10-fold cross validation endowed accuracy of 100%. The consequences indicated that BF Tree algorithms had good possible in automation of LCC of dirt survey data, that in coil, will aid to develop decision prop arrangement to counsel suitable field use arrangement and dirt and water conservation practices.

DursunDelen et al., 2013 [14] In this paper Ascertaining the stable presentation employing a set of commercial measures/ratios has been an interesting and challenging setback for countless researchers and practitioners. Identification of factors (i.e., commercial measures/ ratios) that can precisely forecast the stable presentation is of outstanding attention to each decision maker. In this discover, they retained a two-step analysis methodology: early, employing exploratory factor analysis (EFA) they recognized (and validated) underlying dimensions of the commercial ratios, pursued by employing predictive modeling methods to notice the possible connections amid the stable presentation and commercial ratios.

KalpeshAdhatrao et al., 2009 [15] In this paper an educational association needs to have an approximate prior vision of enrolled students to forecast their presentation in upcoming academics. This helps them to recognize enthusing students and additionally provides them an opportunity to wage attention to and enhance those who should plausibly become lower grades. As a resolution, they have industrialized an arrangement that can forecast the presentation of students from their preceding presentations employing thoughts of data excavating methods below Classification. They have analyzed the data set encompassing data concerning students, such as gender, marks scored in the board examinations of classes X and XII, marks and locale in entrance examinations and aftermath in early year of the preceding batch of students. By requesting the ID3 (Iterative Dichotomiser 3) and C4.5 association algorithms on this data, they have forecasted the finished and individual presentation of newly confessed students in upcoming examinations.

DelveenLuqmanAbd et al., 2013 [16] In this paper an analogy amid three classification's algorithms were learned, these are (K- Nearest Neighbor classifier, Decision tree and Bayesian network) algorithms. The paper clarifies the strength and accuracy of every single algorithm for association in word of presentation efficiency and period intricacy required. For ideal validation patriotic, twenty-four-month data analysis is led on a mock-up basis.

Michal Wozniak et al., 2014 [17] In this paper, a present focus of intense analysis in outline association is the combination of countless classifier arrangements, that can be crafted pursuing whichever the alike or disparate models and/or datasets constructing approaches. These arrangements present data mixture of association decisions at disparate levels vanquishing limitations of established ways established on solitary classifiers. This paper presents an up-to date survey on several classifier arrangement (MCS) from the point of think of Hybrid Intelligent Systems. The article debates main subjects, such as diversity and decision mixture methods, bestowing a vision of the spectrum of requests that are presently being developed.

Brijain R. Patel et al., 2014 [18] In this paper Data excavating is the procedure of discovering or removing new outlines from colossal data sets including methods from statistics and manmade intelligence. Association and forecast are the methods utilized to make out vital data classes and forecast probable trend .The Decision Tree is a vital association method in data excavating classification. It is usually utilized in marketing, surveillance, fraud detection, logical discovery. As the classical algorithm of the decision tree ID3, C4.5, C5.0 algorithms have the merits of elevated categorizing speed, forceful discovering skill and easy construction. Though, these algorithms are additionally unsatisfactory in useful application. After employing it to categorize, there does exists the setback of inclining to select attribute that have extra benefits, and ignoring qualities that have less values. This paper provides focus on the assorted algorithms of Decision tree their characteristic, trials, supremacy and disadvantage.

Gilad Katzet al., 2014 [19] In this paper Decision trees have three main disadvantages: reduced performance when the training set is small; rigid decision criteria; and the fact that a single "uncharacteristic" attribute might "derail" the classification process. In this paper they present ConfDTree

(Confidence-Based Decision Tree) | a post-processing method that enables decision trees to better classify outlier instances. This method, which can be applied to any decision tree algorithm, uses easy-to-implement statistical methods (confidence intervals and two-proportion tests) in order to identify hard-to-classify instances and to propose alternative routes. The experimental study indicates that the proposed post-processing method consistently and significantly improves the predictive performance of decision trees, particularly for small, imbalanced or multi-class datasets in which an average improvement of 5%»9% in the AUC performance is reported.

M.Rajyalakshmi et al., 2014 [20] In this paper, they developed a method for decision tree classification on spatial data streams by means of a data structure called Peano Count Tree(P-Tree). It gives lossless compressed illustration of spatial data set and enables efficient classification and other data mining techniques. Fast calculations of measurements are achieved using P-tree structure. They incorporated Probability Maximization divergence into the partitioning and density-based clustering using PM divergence. They also proposed a new condition of producing terminal nodes so that the decision tree is optimized. The speed is also improved. Experimental results gave better results approx.>96% with different dataset.

V. Conclusion and Future Works

Decision tree is a tree formed data structure that verifies divide and rule approach. Decision tree is used for supervised learning. It is a tree structured model in which the local region is found recursively, with a set of division in a few steps. Decision tree consists of inner decision node and outer leaf. In Future we will work on avariationof Decision trees where classification error will be minimized using Reduced Error Pruning, this algorithm will be based on the principle of calculating the information gain with entropy and reducing the error arising from variance. With the help of this method, complexity of decision tree model can decreased by and the error arising from variance is reduced.

References

[1]. Rodrigo Coelho Barros,,Marcio Porto Basgalupp, A. C. P. L. F. De Carvalho, and Alex AlvesFreitas. "A survey of evolutionary algorithms for decision-tree induction." *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on* 42, no. 3 (2012): 291-312.

[2]. W. Nor Haizan W. Mohamed,,MohdNajibMohdSalleh, and Abdul Halim Omar. "A comparative study of reduced error pruning method in decision tree algorithms." In *Control System, Computing and Engineering (ICCSCE), 2012 IEEE International Conference on*, pp. 392-397. IEEE, 2012.

[3]. Tina R. Patil, and M. S. Sherekar. "Performance analysis of Naive Bayes and J48 classification algorithm for data classification." *Int J ComputSciAppl* 6 (2013): 256-261.

[4]. Smith Tsang, Ben Kao, Kevin Y. Yip, Wai-Shing Ho, and Sau Dan Lee. "Decision trees for uncertain data." Knowledge and Data Engineering, IEEE Transactions on 23, no. 1 (2011): 64-78.

[5]. Rodrigo C. Barros, Márcio P. Basgalupp, André CPLF de Carvalho, and Alex A. Freitas. "Towards the automatic design of decision tree induction algorithms." In Proceedings of the 13th annual conference companion on Genetic and evolutionary computation, pp. 567-574. ACM, 2011.

[6]. Raj Kumar and Rajesh Verma. "Classification algorithms for data mining: A survey." International Journal of Innovations in Engineering and Technology (IJIET) 1, no. 2 (2012): 7-14.

[7]. A.S. Galathiya, A. P. Ganatra, and C. K. Bhensdadia. "Improved Decision Tree Induction Algorithm with Feature Selection, Cross Validation, Model Complexity and Reduced Error Pruning." International Journal of Computer Science and Information Technologies 3, no. 2 (2012): 3427-3431.

[8]. Susan Lomax and Sunil Vadera. "A survey of cost-sensitive decision tree induction algorithms." ACM Computing Surveys (CSUR) 45, no. 2 (2013): 16.

[9]. Mohammed Abdul Khaleel, Sateesh Kumar Pradham, and G. N. Dash. "A survey of data mining techniques on medical data for finding locally frequent diseases." Int. J. Adv. Res. Comput. Sci. Softw. Eng 3, no. 8 (2013).

[10]. AnujaPriyama, Rahul GuptaaAbhijeeta, AnjuRatheeb, and SaurabhSrivastavab. "Comparative Analysis of Decision Tree Classification Algorithms." International Journal of Current Engineering and Technology 3, no. 2 (2013): 866-883.

[11]. Richa Sharma, AniruddhaGhosh, and P. K. Joshi. "Decision tree approach for classification of remotely sensed satellite data using open source support." Journal of Field System Science 122, no. 5 (2013): 1237-1247.

[12]. LeszekRutkowski, Lena Pietruczuk, PiotrDuda, and MaciejJaworski. "Decision trees for mining data streams based on the McDiarmid's bound." Knowledge and Data Engineering, IEEE Transactions on 25, no. 6 (2013): 1272-1279.

[13]. Nirmal Kumar, G. P. Reddy, and S. Chatterji. "Evaluation of Best First Decision Tree on Categorical Soil Survey Data for Land Capability Classification." International Journal of Computer Applications 72, no. 4 (2013).

[14]. DursunDelen, CemilKuzey, and Ali Uyar. "Measuring firm performance using financial ratios: A decision tree approach." Expert Systems with Applications 40, no. 10 (2013): 3970-3983.

[15]. KalpeshAdhatrao, AdityaGaykar, AmirajDhawan, RohitJha, and VipulHonrao. "Predicting Students' Performance using ID3 and C4. 5 Classification Algorithms." arXiv preprint arXiv:1310.2071 (2013).

[16]. DelveenLuqmanAbd, AL-Nabi, , and ShereenShukri Ahmed. "Survey on Classification Algorithms for Data Mining:(Comparison and Evaluation)." Computer Engineering and Intelligent Systems 4, no. 8 (2013): 18-24.

[17]. Michal Wozniak, Manuel Graña, and Emilio Corchado. "A survey of multiple classifier systems as hybrid systems." Information Fusion 16 (2014): 3-17.

[18]. Brijain R. Patel, and Kaushik K. Rana. "Use of Renyi Entropy Calculation Method for ID3 Algorithm for Decision tree Generation in Data Mining." International Journal 2, no. 5 (2014).

[19]. Katz, Gilad, AsafShabtai, LiorRokach, and NirOfek. "ConfDTree: A Statistical Method for Improving Decision Trees." *Journal of Computer Science and Technology* 29, no. 3 (2014): 392-407.

[20]. M.Rajyalakshmi, P.Srinivasulu."High Speed Improved Decision Tree for Mining Streaming Data."International Journal of Engineering Trends and Technology (IJETT) – Volume 18 Number 8 – Dec 2014.