RESEARCH REVIEW JOURNALS
Volume-04
Issue-02
February-2019

ISSN: 2455-3085 (Online)
RESEARCH REVIEW International Journal of Multidisciplinary
www.rrjournals.com[UGC Listed Journal]

A Study on Multi-Layer Classification Using Data Mining Techniques

[1]Pramod Kumar & [2]Dr. Yashpal Singh

[1]Research Scholar, OPJS University, Churu, Rajasthan (India)
[2]Supervisor, OPJS University, Churu, Rajasthan (India)

ARTICLE DETAILS

Article History
Published Online: 20 February 2019

Keywords
Classification Techniques, Decision tree, Multi-Layer perceptron, Neural Networks

ABSTRACT

The measure of information put away in instructive database expanding quickly. These databases contain the valuable data. Information mining is utilized to ponder the information accessible in the instructive field and draw out the concealed learning from it. Factors other than knowledge which influences the scholastic execution of the understudies were examined in this investigation. The examination uncovered that the Multi-Layer Perceptron is more exact than alternate calculations. Our model will Predicts their imprints ahead of time to make better move to improve their standard to get more checks. Arrangement is the way toward finding a model that portrays and recognizes information classes or ideas. Arrangement techniques can deal with both numerical and straight out traits. Building quick and exact classifiers for huge informational indexes is an essential assignment in information mining and learning revelation. Order predicts absolute class layers and arranges information dependent on the preparation set. Characterization is two stages forms. In this paper we present an investigation of different information mining order systems like Decision Tree, K Nearest Neighbor, Support Vector Machines, Naive Bayesian Classifiers, and Neural Networks.

1. Introduction

Data Mining is characterized as removing the data from the immense arrangement of information. Data Mining is a rising control, worried about creating strategies for investigating the interesting kinds of information that originate from instructive settings, and utilizing those techniques to more readily comprehend understudies. Customary issues, for example, enrolment the board and arrangement openings rouse the advanced education organizations to look for better arrangements. The motivation behind information accumulation is to get data to keep on record, to settle on choices about critical issues, to pass data on to other people. We present the outcomes dependent on attributes, for example, adaptability, precision to distinguish their qualities in a Data Mining instrument.

Classification used two steps in the first step a model is constructed based on some training data set, in seconds step the model is used to classify a unknown tuple into a class layer.

Step 1 - Construction of a model

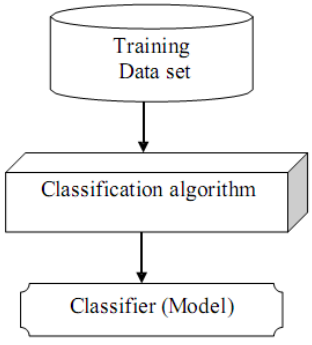

Fig.1 - Model construction step

Step 2 - Model used for unknown tuple

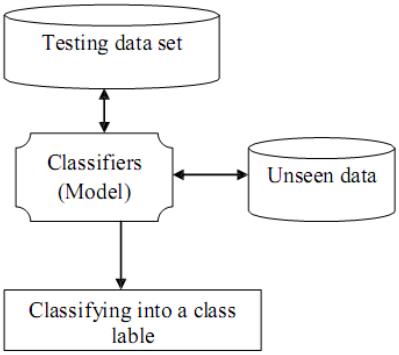

Fig.2 - Use of classifier

2. Literature Survey

In 2012 Akhil jabbar et al. proposed "Coronary illness Prediction System utilizing Associative Classification and Genetic Algorithm". They proposed effective acquainted order calculation utilizing hereditary methodology for coronary illness expectation. The fundamental preferred standpoint of hereditary calculation is the revelation of abnormal state expectation decides is that the found principles are profoundly intelligible, having high prescient exactness and of high intriguing quality qualities. The proposed strategy helps in the best forecast of coronary illness which even helps specialists in their finding choices [1].

In 2013 Akhil Jabbar et al. proposed "Grouping of Heart Disease utilizing Artificial Neural Network and Feature Subset Selection". They proposed another component choice technique utilizing ANN for coronary illness characterization. For rank the characteristics which contribute more towards order of coronary illness they connected distinctive element

choice strategies, and by implication diminish the no. of analysis tests to be taken by a patient. The proposed strategy wipes out pointless and distortive information [2].

In 2014 N. S. Nithya et al. proposed "Addition proportion based fluffy weighted affiliation rule digging classifier for medicinal symptomatic interface". They demonstrated that before model dependent on data gain and fluffy affiliation rule digging calculation for extricating both affiliation standards and enrollment capacities are not achievable. They utilized vast number of unmistakable qualities. They alter gain proportion based fluffy weighted affiliation rule mining and improve the classifier accuracy[3].

In 2015 S. Olalekan Akinola, O. Jephthar Oyabugbe proposed "Exactnesses and Training Times of Data Mining Classification Algorithms: An Empirical Comparative Study". They proposed ponder was intended to decide how information mining characterization calculation perform with increment in info information sizes. They utilized three information mining characterization calculations Decision Tree, Multi-Layer Perceptron (MLP) Neural Network and Naïve Bayes were exposed to changing recreated information sizes. The time taken by the calculations for trainings and exactnesses of their groupings were broke down for the diverse information sizes [4].

In 2015 Jaimini Majali, Rishikesh Niranjan and Vinamra Phatak proposed "Information Mining Techniques for Diagnosis and Prognosis of Cancer". They utilized information digging methods for analysis and anticipation of malignant growth. They introduced a framework for analysis and anticipation of malignant growth utilizing Classification and Association approach in Data Mining. They utilized FP calculation in Association Rule Mining to finish up the examples much of the time found in kindhearted and dangerous patients [5]

In 2016 Nikhil N. Salvithal and R.B. Kulkarni proposed "Examination Management System utilizing Data Mining Classification Technique". The proposed arranged classifier calculations connected on Talent dataset to detect the ability set to pass judgment on the execution of the person. At last depending on precision one most appropriate classifier is picked this strategy has been utilized to develop order tenets to foresee the potential ability that for advancement or not[6].

In 2016 Tanvi Sharma and Anand Sharma proposed "Execution Analysis of Data Mining Classification Techniques on Public Health Care". The proposed examination concentrated on the use of different information mining characterization methods utilizing distinctive AI devices, for example, WEKA and Rapid excavator over the open medicinal services dataset for breaking down the social insurance framework. The level of precision of each connected information mining characterization procedure is utilized as a standard for exhibition measure. The best system for specific informational index is picked dependent on most astounding exactness [7].

3. Characteristics of classifiers

Every single classifier has some quality which differential the classifier structure other. The properties are known as attributes of the classifiers. These qualities are

Correctness: -How a classifier groups tuple precisely depends on these qualities. To check exactness there are some numerical qualities dependent on number of tuple arrange effectively and number of tuple order off-base.

Time: -What amount of time is required to develop the model? This additionally incorporates an opportunity to use by the model to arrange at that point number of tuple (forecast time). In other word this alludes to the computational expenses.

Strength: -capacity to arrange a tuple effectively even tuple has a commotion. Commotion cannot be right esteem or missing worth.

Data Size: -Classifiers ought to be free structure the span of the database. Model ought to be versatile. The execution of the model isn't reliant on the measure of the database.

Extendibility: -Some new component can be included at whatever point required. This component is hard to actualize.

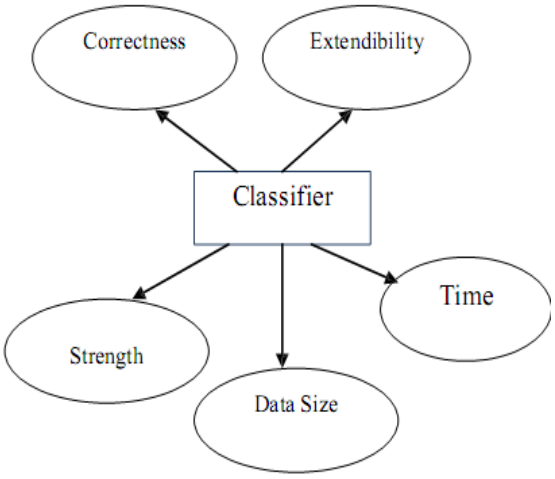

Fig.3 - Characteristic of a Classifier

4. Multi-Layer Classification

In single layer problems, set of instances is D, set of layers is L. For each instance d $\in$ D, select layer set l $\in$ L. So the Single layer representation is (d, l). [1, 2] In multi-Layer problems, set of instances is D, set of layers is L. For each instance d $\in$ D, select layer sub set S $\subseteq$ L. So the Multi-layer representation: (d, S). [1, 2] There are mainly two methods for multi-layer classification problems:

(1) Problem transformation method and
(2) Algorithm adaptation method.

Issue change technique exchanges multi-layer issues into single layer issues. Also, calculation adjustment strategy stretches out explicit learning calculation to deal with multi-layer issues.

Table I: Example of Multi-Layer Problem

Attributes		Class Labels				
A	B	rec	sport	swim	auto	run
A	1	✓	✓	✓		
A	2	✓	✓			✓
A	2	✓	✓	✓		✓
B	1	✓	✓			
B	2	✓			✓	

5. Advantage and disadvantage

Each and every model has some advantage and disadvantage. We give some advantage and disadvantage of these methods

Model	Advantage	Disadvantage
Decision Trees	Easy to interpret and explain.	Do not work best for uncorrelated variables.
K-Nearest Neighbor	Effective if training data is large.	Need to determine values of parameter
Support Vector Machines	Useful for non-linearly separable data	
Naive Bayesian Classifiers	Handles real and discrete data.	Assumption is independence of features
Neural Networks	It is a non-parametric method.	Extracting the knowledge (weights in ANN) is very difficult

6. Conclusion

There are a few characterization strategies in information mining and every single strategy has its favorable position and hindrance. Choice tree classifiers, Bayesian classifiers, order by back spread, bolster vector machines, these systems are anxious students they use preparing tuples to build a speculation demonstrate. Multi-Layer perceptron gives more arrangement precision for anticipating the understudy's evaluation. For future work, the investigation can be reached out with progressively unmistakable ascribes to get increasingly exact outcomes, valuable to improve the understudies learning results. Likewise, examinations should be possible utilizing other information mining calculations to get a more extensive methodology, and increasingly important and precise yields. Some unique programming might be used while in the meantime different variables will be utilized.

References

1. M. Akhil jabbar & Dr. Priti Chandrab “Heart Disease Prediction System using Associative Classification and Genetic Algorithm” International Conference on Emerging Trends in Electrical, Electronics and Communication Technologies-ICECIT, 2012.
2. M. Akhil Jabbar, B.L Deekshatulu & Priti Chandra “Classification of Heart Disease using Artificial Neural Network and Feature Subset Selection” Global Journal of Computer Science and Technology Neural & Artificial Intelligence Volume 13 Issue 3 Version 1.0 Year 2013 International Research Journal Publisher: Global Journals Inc. (USA)
3. N S Nithyaand K Duraiswamy “Gain ratio based fuzzy weighted association rule mining classifier for medical diagnostic interface” Sadhana¯ Vol. 39, Part 1, February 2014, pp. 39–52. Indian Academy of Sciences
4. S. Olalekan Akinola, O. Jephthar Oyabugbe Accuracies and Training Times of Data Mining Classification Algorithms: An Empirical Comparative Study” Journal of Software Engineering and Applications, 2015, 8, 470-477 Published Online September 2015 in SciRes. http://www.scirp.org/journal/jsea
5. Jaimini Majali, Rishikesh & Niranjan, Vinamra Phatak “Data Mining Techniques For Diagnosis And Prognosis Of Cancer” International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 3, March 2015
6. Nikhil N. Salvithal“ Appraisal Management System using Data mining “International Journal of Computer Applications (0975 – 8887) Volume 135 – No.12, February 2016

7. Tanvi Sharma, Anand Sharma & Vibhakar Mansotra "Performance Analysis of Data Mining Classification Techniques on Public Health Care Data" International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization) Vol. 4, Issue 6, June 2016
8. B Rosiline Jeetha "EFFICIENT CLASSIFICATION METHOD FOR LARGE DATASET BY ASSIGNING THE KEY VALUE IN CLUSTERING"International Journal of Computer Science and Mobile Computing A Monthly Journal of Computer Science and Information Technology ISSN 2320–088X IJCSMC, Vol. 3, Issue. 1, January 2014, pg.319 – 324
9. Divya Tomar and Sonali Agarwal " A survey on Data Mining approaches for Healthcare" international Journal of Bio-Science and Bio-Technology Vol.5, No.5 (2013), pp. 241-266 http://dx.doi.org/10.14257/ijbsbt.2013.5.5.25
10. V.Krishnaiah , Dr.G.Narsimha, Dr.N.Subhash Chandra "Diagnosis of Lung Cancer Prediction System Using Data Mining Classification Techniques"(IJCSIT)International Journal of Computer Science and Information Technologies, Vol. 4 (1) , 2013, 39 - 45

