

Analytical Study of Data Mining Predictive Techniques For Health Care

Pooja Singh, Dr. G. N. Singh, Dr. Arvind Singh

Department of Computer Science, Sudarshan College, Lalgaoon, Awadhesh Pratap Singh University, Rewa
Madhya Pradesh, India

ABSTRACT

Article Info

Volume 8, Issue 2

Page Number : 490-493

Publication Issue :

March-April-2022

Article History

Accepted: 10 April 2022

Published: 22 April 2022

Data mining techniques or algorithm is the way which is used to perform the automated operation to take decision based on findings. Algorithm is a procedure or step-by-step logical solution for solving data mining problem, algorithm takes some inputs as a parameter and produce some result as a solution of problem. Another way we defined data mining technique as a data mining algorithm. We have studied various algorithms and techniques such as Classification, Clustering, Regression, Artificial Intelligence, Neural Networks, Association Rules, Decision Trees, Genetic Algorithm, Nearest Neighbor method etc. Here in this paper we have given an algorithm based on clustering objects which is useful to predict possibilities of diseases in particular cases. Algorithm gives the result which is depends on diseases table data for prediction. To analysis the algorithm we have taken graph architecture, which gives the pattern of relationships.

Keyword: Data mining technique, Algorithm, Predictive Algorithm, Data mining for Health Care, Association Rule mining, Pattern, Binary pattern

I. INTRODUCTION

The important and major challenges deal with by healthcare organizations is the precept of quality services at cheaper costs. Quality service means diagnosing patients correctly and managing treatments that are effective and full proof. Poor medical and clinical decisions may lead to disastrous results which are therefore unacceptable. Health care organization needs to minimize the treatment cost of clinical tests. Health care organization can find these types of results by employing appropriate by using

various data mining techniques and decision support systems. Health care data are very large, it includes patient data, resource management data, diagnosis data, medicine data and transformed data. Health care organizations have capacity to analyze these types of data. Medical records of millions of patients can be stored in database and data ware house and so that data mining techniques may help in answering several important and critical questions related to health care system [2].

Data mining is a way or techniques of searching or discovering something new knowledge from data

ware house or from large amount of data by using new techniques and patterns [1]. For this reason certain data mining algorithm can also play important role to find out the searching pattern. We can think Data mining as a Gold mining, and Coal mining where you deal with Gold and Coal. In data mining we/you needs to deals with large amount of data for fetching some important intellectual knowledge for further requirements. Data mining sometimes also called knowledge mining through pattern. In another word data mining is searching the unknown data from the databases or data ware house. Different methods are already available in the process of Data mining but some another new techniques are being developed for this purpose.

Data mining technique algorithm defines the way by which all above process can be easily perform and automated.

1. DATA MINING TECHNIQUES AND ALGORITHMS

Algorithm is a well-defined procedure or step-by-step logical solution for solving data mining problem, algorithm takes some inputs as a parameter and produce some result as a solution of problem. Result will be based on Data mining techniques which are also known as a data mining algorithm. Various algorithms and techniques such as Classification, Clustering, Regression, Artificial Intelligence, Neural Networks, Association Rules, Decision Trees, Genetic Algorithm, Nearest Neighbor method etc., are used for knowledge discovery from databases [1].

2.1. Classification

Classification means it is used to predict Categorical class labels and classifies data based on some sets such as training set and class labels and it can also be used for classifying newly available data [1]. In classification technique, several machine learning algorithms can be used such that Naive Bayes classifier, Support Vector Machine, Linear Regression,

Decision Tree, K-means clustering, Logistic Regression, Artificial Neural Networks etc.[3]

2.2. Clustering

Cluster as name suggest it is a collection of set of objects. In other hand Clustering means grouping of a particular set of objects based on their characteristics, aggregating them according to their similarities. Regarding to data mining, this methodology partitions the data implementing a specific join algorithm, most suitable for the desired information analysis [1]. To fulfil the objective of clustering techniques we have Union and Intersection operation.

2.3. Regression

Regression is used to predict the diseases finding through using different techniques. Regression technique can be adapted for predication. Regression analysis can be used to model the relationship between one or more independent variables and dependent variables. Here variable means may be mined diseases. In data mining independent variables are attributes already known and response variables are what we want to predict [4]. Through Response Variable we can find the result of our mining experiment pattern which is the explanatory variable is modified according to needs. Response Variable is defined as the Dependent Variable, the Outcome Variable, result variable or experimental variable.

2.4. Neural Network

Neural network based on the concept of our working of brain and brain structure. Neural networks are strongly able to adapt to changing input so the network produces the best possible result without the need to redesign the output criteria. Neural network is a set of connected input/output units and each connection has a weight present with it [1].

2.5. Association rule

Association means finding relationship of set of objects that how they are related to each other and

what is probability fulfil the particular case. Another word we can say that Association rule mining means to find frequent patterns, correlations, associations, or causal structures from data sets found in different types of databases such as relational databases, transactional databases, and other forms of data repositories [1].

2.6. Genetic Algorithm

Mainly Genetic Algorithm is based and used as a problem solving technique in order to provide with optimal solution. This is the best way to solve the problem for which little is known [5].

3. ANALYSIS OF DEFFERENT DATA MINIG PREDICTIVE ALGORITHMMS

Using clustering model of algorithm we can classify different diseases in health care. In clustering model collects diseases sample data from different-different groups and separate them by using nested loop of algorithm with their similarity of characteristics or attributes.

Here are the lists of different diseases by which we are doing clustering for prediction of diseases.

| | 1 | 2 | 3 | 4 |
|---|---------|-------|----------|----------------|
| | Malaria | Heart | Covid-19 | Cold and cough |
| 1 | 1 | 0 | 1 | 1 |
| 2 | 0 | 1 | 1 | 0 |
| 3 | 0 | 1 | 0 | 1 |
| 4 | 1 | 1 | 0 | 1 |

Table-1: Diseases and their possibilities

$$A=(1-1,1-3,1-4) \rightarrow 03$$

$$B=(2-2,2-3) \rightarrow 02$$

$$C=(3-2,3-4) \rightarrow 02$$

$$D=(4-1,4-2,4-4) \rightarrow 03$$

The above pairs give possibilities of diseases in a particular condition given in the above table-1. Now

we are going to see the algorithm for the above table-1.

Algorithm

```

int i, j;
n=4;
for(i=1; i<=n;i++)
{
    for(j=1;j<=i;j++)
    {
        if(i[i] == j[j])
        {
            if(i[i]j[j]= =1)
            {
                Printf("%d", i[i]j[j])
            }
            else
            {
                Printf("%d", i[i]j[j])
            }
        }
    }
}
    
```

The above algorithm scans the value of diseases from the table-1 and gives the following binary pattern.

$$A=(1,0,1,1)$$

$$B=(0,1,1,0)$$

$$C=(0,1,0,1)$$

$$D=(1,1,0,1)$$

The set A,B,C,D belongs to the binary set of clustering where set A and set B forms same pattern which means that a particular diseases happens in specific conditions either he follow the set A or set D. Set B and set C follow the opposite pattern and condition from A and D set. Hence we can classify diseases according to the above algorithm's binary pattern and table set pattern.

In regression analysis we are seeing that the set A and set D follows the symmetric relationship that means probability of diseases may happens with both sets conditions.

For the Neural network the above table and binary formation and their sets are connected with its input value and each connection has a weight present with it. So, we can form a graph network by taking above

binary sets. Here are the example of graph and network.

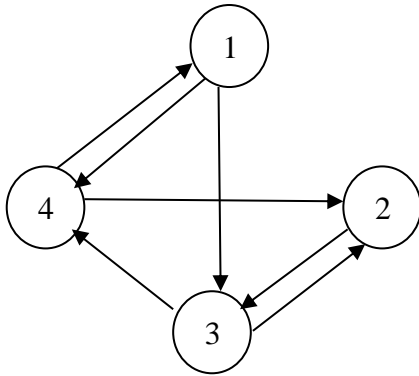


Fig-1: Graph network for connectivity of binary pattern

In the context and analysis of Association rule mining that means that we have to found frequent binary patterns, correlations of between two sets A and D which gives the proper structure and relationship.

II. CONCLUSION

Predictive techniques algorithm defines one of the way by which we tried to find out relationship pattern and structure which predicts about diseases in medical science and health sector. Using this algorithm all processes can be easily perform and automated. The set we have found A, B, C, D forms binary set of clustering where set A and set B forms same pattern which means that a particular diseases happens in specific conditions. Same as regression analysis, neural network, and association rule mining has been discussed in this paper which gives a way to find out some solutions and their analysis.

III. REFERENCES

- [1]. Porkizhi M., "A Study of Data Mining Techniques And Its Applications", IJSART, ISSN [ONLINE]: 2395-1052, Volume 3 Issue 4, APRIL 2017
- [2]. K.Srinivas et al., "Applications of Data Mining Techniques in Healthcare and Prediction of Heart

Attacks", International Journal on Computer Science and Engineering Vol. 02, No. 02, pp. 250-255, 2010

- [3]. Chowdary B V, Radhika Y., "A Survey on Applications of Data Mining Techniques" International Journal of Applied Engineering Research ISSN 0973-4562, Volume 13, pp.5384-5392, Number 7, 2018.
- [4]. Tamilselvi R., Kalaiselvi S., "An Overview of Data Mining Techniques and Applications", International Journal of Science and Research (IJSR), India Online ISSN: 2319-7064, Volume 2 Issue 2, February 2013
- [5]. Jain Nikita, Srivastava Vishal, "DATA MINING TECHNIQUES: A SURVEY PAPER", IJRET: International Journal of Research in Engineering and Technology, Volume: 02 Issue: 11, Nov-2013.

Cite this article as :

Pooja Singh, Dr. G. N. Singh, Dr. Arvind Singh, "Analytical Study of Data Mining Predictive Techniques For Health Care", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 8 Issue 2, pp. 490-493, March-April 2022.

Journal URL : <https://ijsrcseit.com/CSEIT22824102>