International Journal of Scientific & Engineering Research, Volume 5, Issue 6, June-2014
1175
ISSN 2229-5518

A survey on blood transfusion based on data mining techniques

S.Asha rani , Dr.S.Hari Ganesh

Abstract -Data mining is the process of collecting relevant data from enormous amount of data..Requirement of blood is increasing gradually due to accidents, surgeries etc. Blood transfusion play an important role in healthcare. Blood donor prediction provide essential details to medical professionals to increase the number of voluntary blood donors in future. This paper focus on analyzing the Blood transfusion and blood donors behavior using data mining techniques.

Keywords: Data mining, Blood Bank, Blood transfusion dataset, Blood donor, classification algorithms

1.INTRODUCTION

Data mining [1] refers to the process of extracting knowledge from large amounts of data. Data mining that discover the hidden information from large database. Data mining mostly used in all fields such as medical fields, marketing, Image processing etc. Classification, association rules, clustering are some important techniques used in the data mining.

ASSOCIATION RULE:

Association rule is used to identify relationship or simply identify association between the variables in large database. It is the most important concept that used in data mining .In this association rule , Apriori algorithm, FP growth are some important algorithms used to discover relationship between variables.

CLASSIFICATION :

Classification that classifies data based on training set and values. Classification techniques that used to split the records into small segment. These small segments are easy to identify and discover the hidden information. In this classification techniques decision tree play an important role .Bayesian classification, classification by Back propagation are some other important concept used in classification.

CLUSTERING:

Clustering is the process of grouping similar objects into one group called cluster. Clustering used in many fields like information retrieval, Image analysis, machine learning etc. Hierarchical clustering, centroid based clustering, distribution based clustering are some important clustering techniques used in data mining.

BLOOD DONATION:

Blood donation play an important role in saving life.. It is the basic need for all treatments in medical field. Many developed countries, blood donors are mostly regular and volunteer donors and new volunteer donors are also increased gradually day by day. In developing countries, blood donors are paid and not volunteer .They are mostly donate their blood when their relations or friends are in need of blood.

2. RELATED WORK

Arvind Sharma [C], in his research work predict the number of blood donor through their age and blood group. In this paper, real-time dataset are implemented and measure the accuracy of blood donor prediction using classification algorithm. WEKA tool used for this prediction.

Vikram singh [D] used the Blood transfusion dataset from UCI machine learning lab. They explain interactive KDD process performed in short time and less effort with

interaction between human and computer system.

Shyam sundaram [E] in his research work used Blood transfusion dataset from UCI repository. CART algorithm used for testing in WEKA tool.DB2K7 (Donated Blood in 2007) and RVD (Regular Voluntary Donor) are two models used for testing. DB2K7 used donated blood in 2007 attribute as nominal and RVD used recency and frequency as the key indicator.RVD show better accuracy than DB2K7.

Alaa hamouda et al [F] The purpose of that research work is to count the RBC automatically using Image processing techniques. Image processing approach such as equalization, segmentation are used. In RBC extraction , techniques such as classification, K-means clustering, hybrid between clustering and classification, learning using decision tree are used. In this testing, learning using decision tree technique shows high accuracy than other techniques.

Shyam sundaram et al.[G] presented a comparison among various location in india about blood donation and to predict regular voluntary donors are from which location in india. Blood transfusion dataset from UCI repository is used. Geo location attribute is added.

Ivana D.Radojevic et al. [H] monitor the water quality by coliforms presented in the reservoirs. total coliforms are analyzed and modeled by using classification and clustering analysis. This analysis shows the dominant bacterial community, number of total coliforms is connected to human activities, morphometric characteristic of reservoirs etc.

Wen-Chen Lee et al.[I] developed the system implementing classification and clustering algorithms to predict the blood donors behaviors and understand their problems and to increase voluntary blood donors. K-means clustering and classification techniques such as Decision tree,Naïve bayes,NB tree are used randomly with blood transfusion dataset of UCI repository for demonstration purpose**.**

3. SUMMARY:

This paper is used to analyze the blood donation details and blood donors behavior such as there are voluntary blood donors are not. Various authors explain different techniques and algorithms for prediction of blood donors. This analyzes provide solution to increase voluntary blood donors rate especially in developing countries like India and awareness about blood donation among people.

4. REFERENCES

[A] Jiawei Han, Micheline Kamber and Jian pei.: "Data Mining; Concepts and Techniques", Morgan Kaufmann Publishers. March 2006. ISBN 1-55860-901-6

[B]. Pieter Adrians, Dolf Zantinge: "Data Mining", Addison Wesley, 2000

[C] Arvind Sharma and Gupta.P.G, "Predicting the number of blood donors through their age and blood group by using data mining tool", International journal of communication and computer technologies (IJCOTS), vol-1, and issue: 02, September 2012.

[D] Vikram Singh and Sapna Nagpal, "Interactive knowledge discovery in blood transfusion data set", VSRD international journal of computer science and information technology (VSRD-IJCSIT, vol-1(8), 2011, 541-517.

[E] Santhanam. T and Shyam sundaram, "Application of CART algorithm in blood donors classification", Journal of Computer science 6(5), 548-552, 2010.

[F] Alaa Hamouda, Ahmed Y.Khedr and Rabie A.Ramadan ," Automated Red Blood Cell Counting", International journal of computing science,vol.1,No.2,february 2012.

[G] Shyam sundaram and Santhanam.T, "A comparison of blood donors classification data mining models", Journal of theoretical and applied information technology, vol 30 No.2,
31 august2011

[H] Ivana D.Radojevic et al, "Total coliforms and data mining as a tool in water quality monitoring", African journal of microbiology research, vol.6 (10), 16 march 2012

[I] Wen-chanLee and Bor-wen cheng, "An intelligent system for improving performance of blood donation", Journal of quality vol.18, Issue No.11, 2011.