Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 46 (2015) 682 – 688

Procedia Computer Science

International Conference on Information and Communication Technologies (ICICT 2014)

Automated Prediction of RCT (Root Canal Treatment) Using Data Mining Techniques: ICT in Health Care

Kavita Choudhary[a,*], Pinki Bajaj[a]

[a]*Assoc. Professor, Department of CSE & IT, ITM University, Gurgaon,122017, India*
[a]*Student, Department of CSE & IT, ITM University, Gurgaon,122017, India*

Abstract

With the use of various data mining techniques to be implemented on rapid minor tool, diseases can be detected .This paper makes use of classification technique based on cross validation and decision tree to detect RCT before seeking advice from the doctor. The main aim of this paper is to diagnose whether there is a need of RCT or not with different input values of an attributes. All these symptoms or attributes are analyzed to reach a decision that person with an age greater than 18 and having sensitivity problem have greater chances of root canal treatment. These techniques make the doctors and patients to take an early decision.

Peer-review under responsibility of organizing committee of the International Conference on Information and Communication Technologies (ICICT 2014)

Keywords: Cross Validation Algorithm; Decision Tree; Apply Model; KNN Classification; Age; Brushing; Sensitivity; Tooth decay; Diet; Smoking.

1. Introduction

Information and Communication Technology (ICT) is reforming day-to-day life of human and the way of communicating with each other. One of its applications in health sector is e-Health which provides better care with increase in quality of efficiency. It helps in reducing the operating cost with an increment in patient's protection. With healthcare Information Communication Technology (ICT), there is an amendment in health

* Corresponding author. Tel.: +919999667797;
E-mail address:kavitapunia@gmail.com

Peer-review under responsibility of organizing committee of the International Conference on Information and Communication Technologies (ICICT 2014)
doi:10.1016/j.procs.2015.02.124

care process so that information can be in the hands of professionals who can deliver better outcome to the patients[13].Now a day's these technologies have a great impact on the medical field and no one can imagine any treatment without its usage. Instead of taking a manual hypothesis about any disease and continue for the treatment, a thorough understanding of diseases is the most crucial part for any physician as it's a matter of someone's life. By using and adapting the changes occurring in an environment regarding medical domain not only help doctors for making decision but also satisfies the patients. Such technologies are designed to mend the quality of care delivered through earlier detection of any disease which further reduces the cost and saves time. Physicians can gather information with the help of technology which helps in knowing more about any domain. Medical science has also developed itself accordingly with the advancement in the technology. Information and technology helps to store, retrieve and manipulate data and is broadly used in the medical sectors as Health Information Technology (HIT). This technology helps in building a system by networking various computer and communication attributes to move health information. Operating microscopes, Apex locators, ultrasonic are some of the technologies that make the doctor more understandable with the tiny details. With the rise in population and increasing cost of healthcare, there is a great use of health informatics to save money and time. These technologies are not only used in healthcare but also in domains like analyses, education, research and training. This paper explains the requirement of Information technology in Root Canal Treatment[9]. Root canal treatment is a tooth procedure that undergoes some steps and protects the natural teeth from removal and some further infection. During this procedure, the nerve and pulp are removed, the tooth is cleaned and the canals are shaped to seal the filling material. It remains here till the permanent filling or a crown is placed on the top of the tooth[14]. Rapid minor is a tool which provides an integrated environment for data mining, predictive analytics and so on. Various data mining techniques like classification, association rule mining, fuzzy techniques, co clustering are implemented on this tool to get an output which helps in an early detection of any disease. Data mining techniques like cross validation, decision tree and KNN classification are used in this paper to diagnose whether a person opt for root canal treatment or not. In the same way, an apply model is used to get a performance based accuracy. Decision tree is a model that generates a tree like structure from the given set of input values which helps in making a decision. Given input attributes are represented as a path from top to bottom and its values are shown by the leaf node. This paper implements the decision tree using cross validation technique. It is used to measure how accurately a model will work out. Cross validation consists of two sub processes: one is training which performs the work of processing to be put into the second sub process and another one is testing which is used to analyze the performance of the model. In KNN classification, similar inputs of both test and training examples are compared. It detects the nearest neighboring class and is easiest among the other machine learning. Root canal treatment detected by the data mining techniques explained earlier, uses some values of symptoms which help the physician in understanding a problem more deeply so that to begin a treatment earlier and can overcome the further harm or infection. Other techniques can also be used to predict the diseases as early as possible.

2. Literature Survey

Buczak[1] described an approach for building models to Mine Synthetic Electronic Medical Records (EMR) and for performing clustering of data set. The records of patients who exhibit an illness are generated by selecting the most appropriate set of clustering using matrices. Chung[2] proposed a system in the field of bioscience including analysis of diseases, managing information, records of hospital and so on; by applying relational database techniques based on probability. Megala[3] estimates the details of children with Immunization by using Artificial Neural Network and classification technique and deduce that a body may undergo serious health problems due to the large amount of polio virus. Sherafat[4] presented a framework consist of patient's data with the knowledge from different institutes and determined how a decision can be made by combining various algorithms with this framework to work consistently in an environment. Karima[5] proposed a methodology to fragment Data Warehouse (DW) relations which specifies the combination of construction and clustering techniques. Fragments are then allocated according to their frequency to the corresponding site. While comparing with the centralized, Distributed Data Warehouses gives better performance for all the specific queries and the time of execution for different queries has been reduced by

some percentage. Chauhan[6] proposed a framework which includes a relation of OCT and perimeter data so that glaucoma can be diagnose as early as possible. It also predicts the progression of glaucoma and its early detection can prevent early vision loss in the patient. Muangnak[7] examined the learning disabilities in students by using two classification techniques, Naive Bayes Classifier and Decision Tree with Model C4.5. This Study proposed that the Naive Bayes Classifier has less accuracy in the classification decision tree which concludes that gender, age and grade were not affected and disabilities of learning. Lin[8] applied the business model co clustering algorithm (BCA) which provides the connectivity patterns to higher up the strength of connection in human brains where cortical areas exists. This study results that the developed BCA is well suited and consistent both in its speed and stability. Balasubramanian[9] examined the factor of risk in Tamil Nadu State, Krishnagiri District, India with the higher level of fluoride in water. With the use of clustering algorithms, patterns are determined which gives decisions that the more amount of fluoride can cause neuro, dental and skeletal problems. Naseri[10] focused on the amount of filling solution and sealer by using Micro-Computed Tomography to evaluate Obturation Techniques and determined that teeth are gap-free with the highest- lowest volume percentage of filling and cold lateral condensation obturation techniques. Shukla[11] presented the data mining's application in clinical, administrative, educational and research aspects of Health Informatics. Various techniques of Data mining like association rule mining, CART (Classification and Regression Tree), clustering, classification, regression, are broadly used in health sector. Its algorithms improve the prediction, diagnosis and classification of diseases. This paper analyses the techniques and also highlights the importance of locally frequent patterns that are used to discover diseases such as breast cancer, lung cancer, heart diseases and so. After this review, the well performing algorithm of data mining is identified and evaluated. Collste[12] represented a series of prospect from distinct fields including medicine, philosophy and computer science where information and communication technologies (ICT) with respect to medicine have changed the image. This paper is written to collect the content to be present in session of panel where various issues regarding computing and human choice can be discussed as these are the most crucial issues of the individual and society regarding the effect of Information Technology. Rudowski[13] explained how various health care services can be accessible for the citizen at any time and in any place. This is done with the help of e-Health to avail the services where physician may not be available. This paper observed the shift from medical services to prevention and health increment which gives the better quality and safer care. Lsmann[14] discusses the various problems regarding root canal preparation and also described and compared various instruments and techniques to be used in the root canal treatment. This paper concluded that the use of one of the instrument NiTi for bend root canal results in preferable centered preparation but does not give proper cleanliness. While comparing with other instrument, this is superior with active cutting blades and must be use with proper guidelines. Ranjan[15] explains how data mining plays a crucial role in pharmaceutical industry which helps in discovering and mining patterns from the data to observe it. Advancement in the technology helps the pharmaceutical firms in managing and developing new products. It established the ability of data mining to improve the quality of decision making process and concluded that when all the created molecules in an industry grouped together using clustering then it would be easier to find the influencing element belonging to a particular group. Durairaj[16] analyzes the impact of different approaches and techniques on health sector with the use of different tools whose goal is to change the data into knowledgeable information. Its aim is to create a report of health sector containing different techniques of data mining applications and how the complexity of data transaction in health care can be decrease. With all these studies and comparisons among techniques, it has been concluded that prophecy of diseases using data mining increment the accuracy with a decrement in effort of diagnose. It also observed that predicting diseases with more than one technique yields better results. Vijiyarani[17] uses various data mining techniques and explores that how these could be helpful in detecting some diseases like heart disease, diabetes and breast cancer. With this diagnose, it is useful in reducing the number of tests required for detecting a particular disease which further reduces the time. Various attributes for different problems are taken and are implemented using different algorithms of data mining like Decision Tree, classification via clustering, Naïve bayes and Association rule discovery which deduce some result regarding accuracy.

3. Proposed Model

Data samples of various attributes are collected which uses the random values of age, brushing, sensitivity, tooth decay, diet and smoking for different patients. This paper provides the possibility of root canal treatment using different data mining techniques[9]. By using the decision tree and the classification technique for these values, the probability to undergo RCT is detected.

N: Input values for RCT_ANALYSIS

The Attributes are: - BRUSHING {regular, irregular}, AGE, DIET {nutritious, unhealthy}, SMOKING {yes, no}, SENSITIVITY {yes, no}, TOOTH DECAY {yes, no}, RCT {0, 1}

Detection of RCT using Decision Tree and cross validation:-

rct_decision_tree_cross_validation (N, S, O)

Where:

N -Input data (values in each attribute set)

S - Attribute Set

O-Output based on relation between attributes

Attribute Set (S) ←BRUSHING, AGE, DIET, SMOKING, SENSITIVITY, TOOTH DECAY

- For each S (j), where j is the index partition converted into n partitions of input data Nj.
- Every time of n executions with a different training set N-Nj, the result is verified on Nj and for all the values of S (j) branching is done.
- The input data is run by n times on cross validation. Dataset with size (n−1)/n times the size of the original dataset is tested on different training. Repeat until each level with data partition is done.
- The root node of the tree is the attribute with highest index.
- Check that no attribute is left without partitioning

Decision tree t (T)

{

Input data with Attribute values are segmented into n number of folds.

Out of which, one fold is for training date set and the remaining (n-1) folds are for test data sets.

Both the test and the training data sets are compared with each other.

The comparison is done until all training sets are taken into consideration.

IF (training_set_splitting> threshold) condition is true (Condition_Result = 1) then RCT exists

ELSE condition is false (Condition_Result = 0) then RCT does not exist

}

4. Analysis

Fig. 1. gives the view about the process in rapid minor tool with the techniques of data mining. In this, attributes like brushing, age, smoking, diet, tooth decay and sensitivity with different input values are imported to the process under read excel and various stages are connected to each other accordingly. These connections are done to make an efficient use of data mining techniques. This Figure represents the cross validation technique with apply model and decision tree Fig. 2. shows the nearest neighbor model classes of KNN Classification with the number of dimensions. Fig. 3. depicts the performance based accuracy of about 93.06% to detect RCT using cross validation based on decision tree. Performance determined in this figure helps to know how accurately a process is giving the result for diagnoses of root canal treatment. Fig. 4. shows that by comparing various input attribute values of the data set imported in the process, a tree is created by using the attributes as a node and edges to show its range. This figure has taken the entire attributes as symptoms with some ranges to reach an outcome. On behalf of this, a decision is made for persons to undergo RCT. The attribute values 1.0 and 0.0 gives an idea about the root canal treatment to be taken or not.

5. Simulation Results

It consists of simulation result of this paper showing how various data mining techniques are implemented.

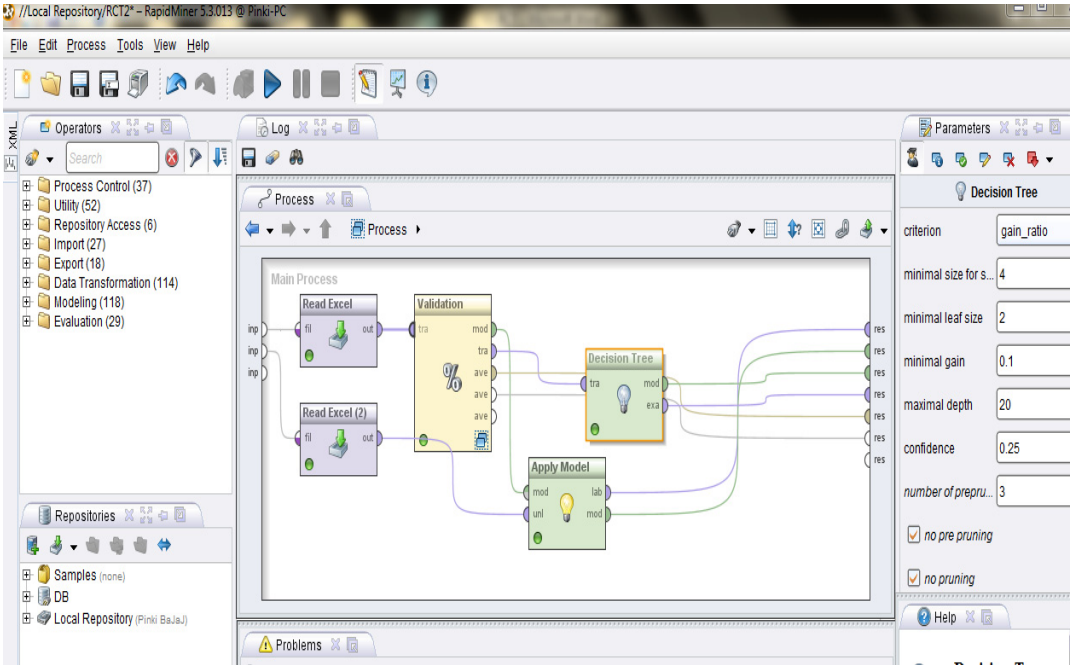

Fig. 1. Model showing cross Validation and decision tree

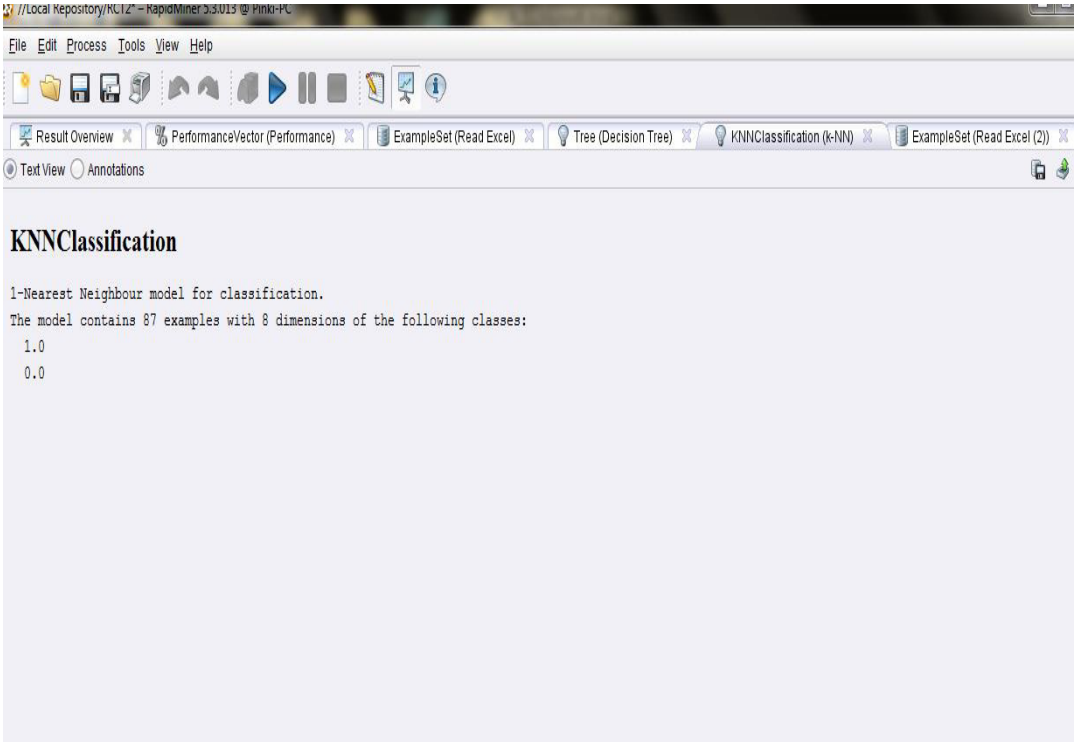

Fig. 2. KNN Classification

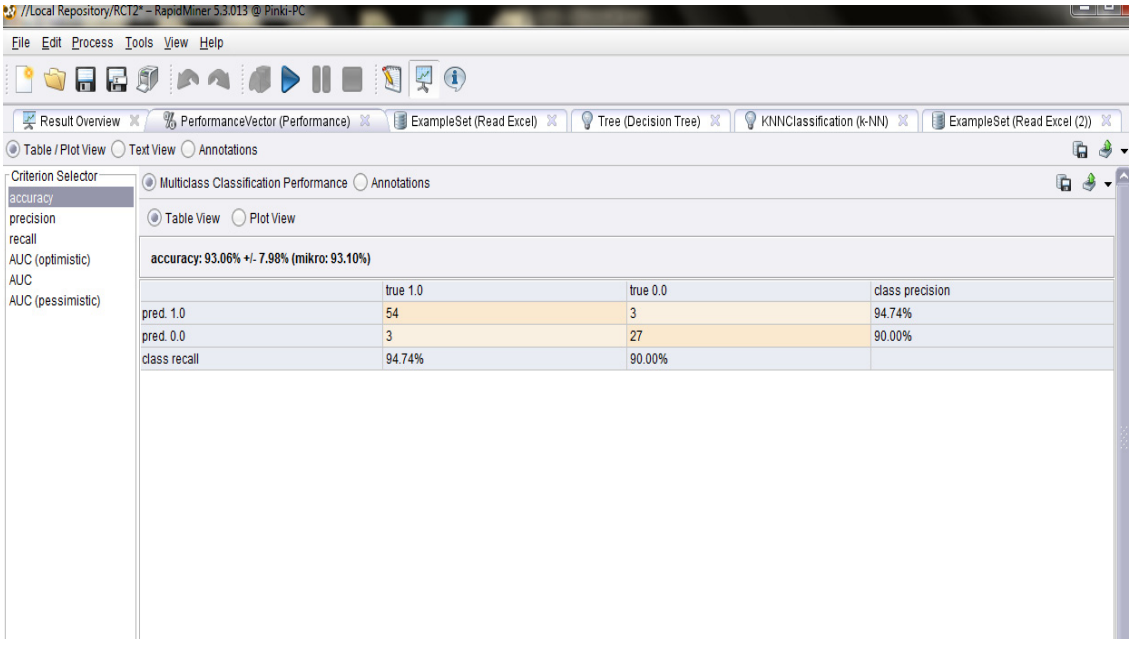

Fig. 3. Accuracy

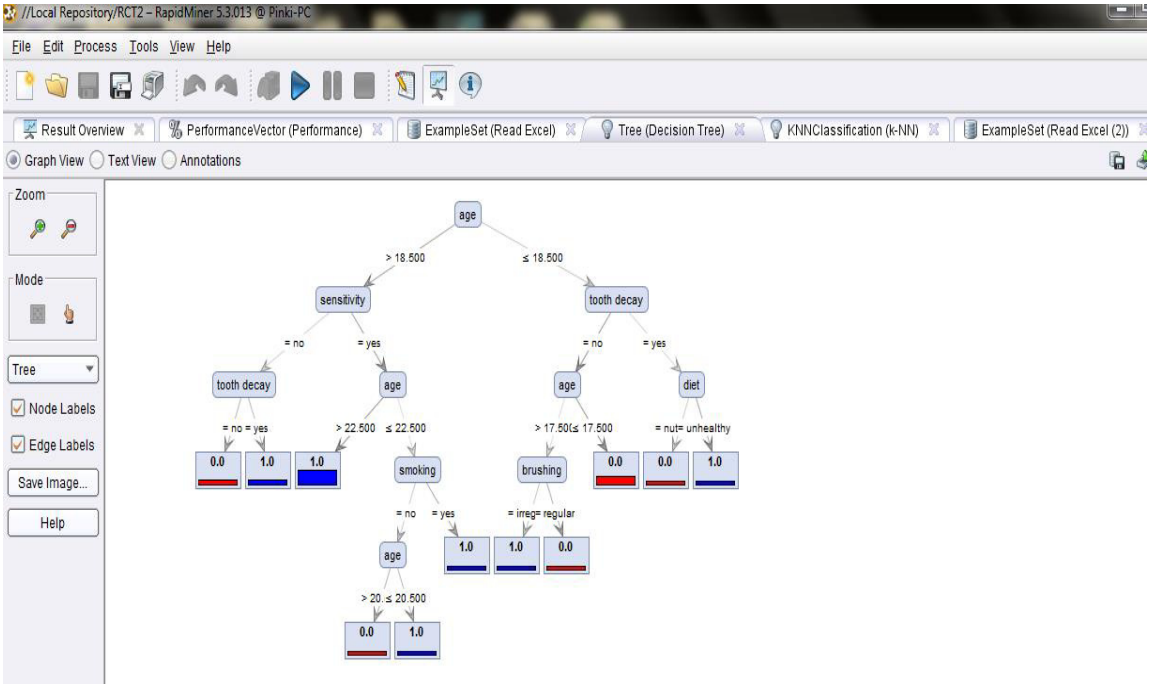

Fig. 4. Decision Tree

6. Conclusion

With this paper, it is concluded that various attributes like BRUSHING, AGE, SMOKING, DIET, SENSITIVITY, TOOTHDECAY with different values undergo a process that results in some prediction of undergoing RCT. As technology is advancing at a great pace which reduces human efforts and time. In the same way with the help of this rapid minor tool which contains data mining techniques, diseases can be detected earlier. This paper evaluated how best possibly RCT can be diagnosed before medical consultation. Using cross validation technique based on decision tree concludes that the persons with age more than 18 and bears sensitivity due to tooth decay problem are having high possibility of root canal treatment. Here, age attribute is more prominent. When input datasets of different patients are applied on rapid minor tool, then the performance comes up with an accuracy of about 93.06%.

References

1. Buczak AL, Moniz LJ, Feighner BH, Lombardo JS. Mining Electronic Medical Records for Patient Care Patterns. *Computational Intelligence and Data Mining* 2009;146-153.
2. Chung Ping Tsai, Afzal F, Hsiao Hsin-Hua. A Software System Development for Probabilistic Relational Database Applications for Biomedical Informatics. *Advanced Information Networking and Applications Workshops* 2009;1002-1007.
3. Dr. Hemalatha M, Megala S. Mining techniques in health care: A survey of immunization. *Journal of Theoretical and Applied Information Technology* 2011;25:2 63-70.
4. Kazemzadeh Reza Sherafat, Sartipi Kamran. Interoperability of Data and Knowledge in Distributed Health care Systems. *Proceedings of the 13th IEEE International Workshop on Software Technology and Engineering Practice* 2005;230-240.
5. Karima Tekaya, Abdellatif Abdelaziz, Ounalle Habib. Data mining based fragmentation technique for distributed data warehouses environment using predicate construction technique. *IEEE Computer Society (NCM)* 2010;63-68.
6. Chauhan Kinjan, Dr. Gulati Ravi. A Proposed Framework for diagnosis of Glaucoma - A Data mining Approach. *International Journal of Engineering Research and Development* 2012;3:56-9.
7. Muangnak N, Pukdee W, Hengsanunkun T. Classification Students with Learning Disabilities Using Naive Bayes Classifier and Decision Tree. *Sixth International Conference on Networked Computing and Advanced Information Management* 2010;189-192.
8. Lin C, Pai D, Lu S, Muzik O, Hua J. Coclustering for Cross-Subject Fiber Tract Analysis Through Diffusion Tensor Imaging. *IEEE Transaction on Information Technology in Biomedicine* 2010;514-525.
9. Balasubramanian T, Umarani R. An Analysis on the Impact of Fluoride in Human Health (Dental) using Clustering Data mining Technique. *International Conference on Pattern Recognition, Informatics and Medical Engineering (PRIME)* 2012;370-375.
10. Naseri Mandana, Kangarlou Ali, Khavid Atefeh, Goodini Mostafa. Evaluation of the Quality of Four Root Canal Filling Techniques. Using Micro-Computed Tomography. *Iranian Association of Endodontists and Iranian Center for Endodontic Research (ICER)* 2013; 8:389-93.
11. Dr. Shukla DP, Patel SB, Sen AK. A Literature Review in Health Informatics Using Data Mining Techniques. *International Journal of Software and Hardware Research in Engineering* 2014.
12. Collste Goran, Duquenoy Penny, George Carlisle, Hedström Karin, Kimppa Kai, Mordini Emilio. ICT in medicine and health care:Assessing social,ethical and legal issues. *International Federation for Information Processing (IFIP)* 2006;223:297-308.
13. Rudowski Robert. Impact of Information and Communication Technologies (ICT) on Health Care.
14. HuLsmann Michael, Peters Ove A, Dummer Paul MH. Mechanical preparation of root canals: shaping goals, techniques and means. 2005;10:130-76.
15. Ranjan Jayanthi. Application of data mining techniques in pharmaceutical industry. *Journal of Theoretical and Applied Information Technology* 2007;61-67.
16. Durairaj M, Ranjani V. Data Mining Applications In Healthcare Sector: A Study. *International Journal of Scientific and Technology Research* 2013;2:1029-35.
17. Vijiyarani S, Sudha S. Disease Prediction in Data Mining Technique – A Survey. *International Journal of Computer Applications and Information Technology* 2013;2:117-21.