ISSN 2395-1621

Clinical Decision Support Systems The Imperatives for Consolidating The Healthcare Informatics

[#1]Tejas Bramhecha, [#2]Pranjali Deshmukh, [#3]Ronak Agrawal, [#4]Parul Khushu

[1]jaintejas511@gmail.com,
[2]pranjalisd@gmail.com,
[3]agrawal.ronak95@gmail.com,
[4]khushu.parul@gmail.com

[#1234]Department of Information Technology,
MIT Academy of Engineering, Pune, India.

ABSTRACT

The top-notch medical research in Clinical decision support systems (CDSS) has truly made a bigger impact on healthcare informatics. It has not only revolutionized the way medical facilities are being serviced to patients but also have helped doctors & physicians to improve the quality & efficiency of diagnosis. This paper has provisioned solution for manifold challenges & is focused on sweeping views of Healthcare industry. The barriers in decision making call into question the use of such systems that can avail the precise predictions at the point of patient care & hence CDSS development in compliance with accuracy is the real dare for developers today. Right from automating processes to analyzing data is actually a complex but an obligatory task for accurate interpretation. We all know that Healthcare organizations are going through technological shifts & data revolution so the intent here is to construct a framework for escalating patients' safety by reducing error rate & adverse effects of medications.

Keywords: CDSS, Machine Learning (ML), predictive healthcare, big data, evidence-based medicine (EBM) etc.

ARTICLE INFO

Article History

Received: 4th May 2017

Received in revised form :

4th May 2017

Accepted: 6th May 2017

Published online :

7th May 2017

I. INTRODUCTION

Nowadays, healthcare informatics plays a distinctive role in our social life. New medicines, drugs & preventive measures are being invented day by day & henceforth the quality & availability of these medicines in the market is a real matter of concern. In today's fast changing world, decision makers are facing growing demand for specialists' opinions in order to comply with patient safety & healthcare quality. Therefore due to prominent relevancy in healthcare informatics, it becomes necessary for us to analyze the available data & predict what could be the next step in order to make sure that it inhibits the medical errors that have been generated at the time of decision making. Clinical decision support system is just an application that let doctors use them for decision making while dealing patients with the timing of use either pre-diagnosis, during diagnosis or post diagnosis.

According to the survey of *Times of India*, In 2011-12 there were 5.2 million people in India & 43 million people across the globe who were injured because of misdiagnosis & uncategorised care [11]. Medical errors like unnecessary treatments, medication mistakes, never events & missed warning signs etc. may reduce the life expectancy of patients suffering from critical diseases like cancer & diabetes so traditional ways of treating a patient won't suffice to quench our requirements when our specialists are grappling with difficulties in making an assessment. *World Health Organization, Europe* has stated certain benefits of such systems. According to them, 750000 harm inflicting errors, 3.2 million days of hospitalization, 260000 cases of permanent disability & 95000 deaths per year can be reduced if European Union adopt systems that use valid, clean, consistent & sound information for strategic decision making [12].

Hereby we aim at building an application that will develop protocols for decision making wherein our objective is to provide an alert & passively responding to physicians' input. The use of Machine learning is so pervasive today that we use it without even knowing it in our daily applications. It applies various techniques & technologies for capturing knowledge from vast dataset &

transforms it into usable information. Therefore, an integration of CDSS with ML will certainly yield us a productive application that incorporates the guidelines to simulate human thinking & help doctors while giving them a secondary opinion over specialists' concern of treating a patient.

The CDSS will help us by providing an insight to power decision making according to the evidence level & number of clinical trials. A working definition has been proposed by Robert Hayward of the Centre for Health Evidence: 'Clinical decision support systems link health observations with health knowledge to influence health choices by clinicians for improved health care' [10]. How best to go about forecasting which treatment must be given to a patient is still a matter of debate but the proposed system will efficiently look for better prognostication using spread drift & risk-free rate. The accuracy of the system can be determined by cross checking the predictive index to the specialists' hand-drawn ground truth.

II. Clinical Decision Support System

A. Concept

Considering current death rate scenario, CDSS is accounted as one of the important strategies to safeguard medical treatments. It has long been known that the prescription, delivery and use of drugs in healthcare are processes prone to errors that may jeopardize patient safety. CDSS processes clinical signs, symptoms & laboratory tests so as to output diagnostic and therapeutic recommendations. It basically focuses on the way knowledge bases are built, and how concluding mechanisms and group decision-making practices are utilized. What makes this application more compelling is it must be integrated with Electronic Health Record (EHR) software. Predicting a treatment seems to be most fascinating idea today. The job of finding patterns from the myriad dataset is extensively being done using ML & is one of the key aspects of it. With so much information available, we can never be sure that we are not overlooking some important facts pointing the way to surpass performance. CDSS flags patients that were not properly diagnosed, missed or given the wrong dosage of medication. It helps us to avoid missing incompatible medications. Identified features of CDSS are as follows:

TABLE I. CDSS FEATURES

Sr.No	Features
1.	Provides patient assessments & treatment recommendations automatically.
2.	Automatically incorporates patient data from EHR.
3.	Links with electronic patient charts to support workflow integration.
4.	Promotes action rather than inaction.
5.	Provides research evidence to justify assessments & recommendations.
6.	Engages local users during system development.
7.	Gives decision support results to patients as well as providers.
8.	Reminders for preventive services including screening for risk factors.
9.	Alerts when indicators indicate any risk factors.

There are two types of CDSS which include knowledge based & non-knowledge based systems. Knowledge-based systems often use if-then rules on the other hand non-knowledge based systems have machine learning rules associated with it [10]. Based upon usage, CDSS systems are classified into three types. Pre-diagnosis CDSS is used to get prediction done before diagnosis. CDSS during diagnosis is used to reconsider diagnostic choices & make them more accurate. Post-diagnosis CDSS is used to mine data to derive connections between patients & their past medical history & clinical research to predict future events.

B. Electronic Health Record (EHR)

The organizational change can be understood by considering the interactions between the content and processes of change within the organization. The organizational entities involved in collecting the patient's data at a hospital are centrally connected to the central database as shown in Fig.1.The data collected through the EHR system is further exposed to the CDSS module. By increasing clinicians' available time for direct patient care, increased application of clinical pathways and guidelines, facilitating the use of up-to-date clinical evidence and patient satisfaction, EHR helps with the smooth functioning of hospital chores. Data is growing faster than medical science can consume it. The life course assessment data of a patient can be segregated at a centralized point so that it can be easily captured at later point whenever is needed.

Due to EHR, the patients will not only get benefitted but the quality of healthcare services will also be improvised & identification of health care deception can be done easily.

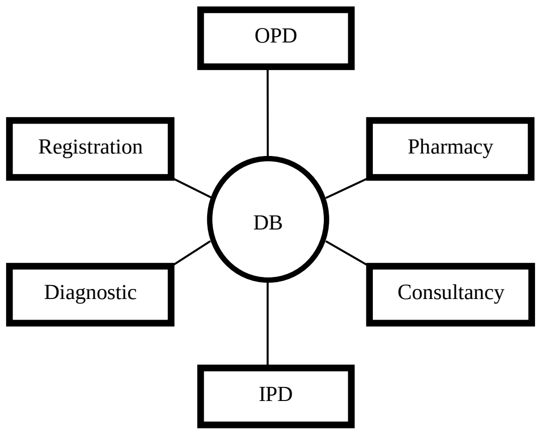

Fig.1. Electronic Health Record

C. Meta-model of proposed system architecture

The workflow of a system along with CDSS is described in Fig.2. Architecture given here explains different data sources, storages & output.

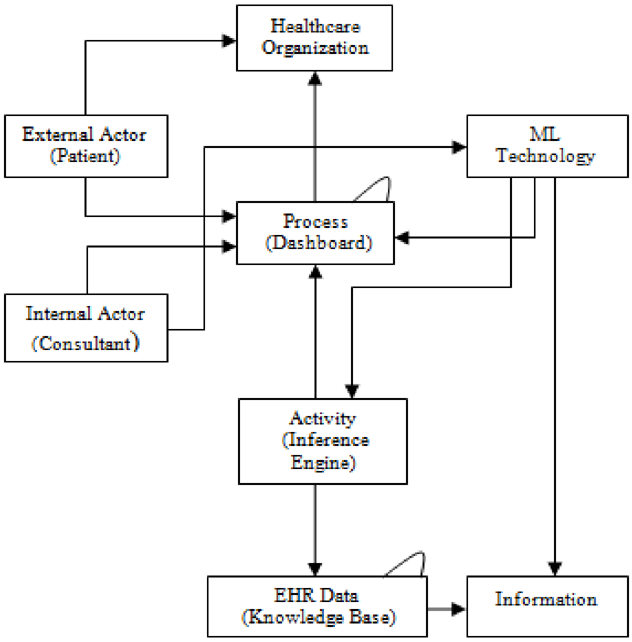

Fig.2. Proposed architecture [16]

The Fig.2 clarifies the relationships of processes, actors, information and technology in the context of Machine Learning. The proposed system functionality is given as below:

1) Trigger at process block will initiate the CDSS workflow. The input data is represented by the components of quality data set.

2) EHR Data will be converted into information & then ML technology block will generate the inferences and hypothesis based on information available. The Hypothesis may include prediction, log, display alerts, reminders etc.

3) Healthcare organization will be associated with process block which deals with a user interface where medication, precautions & preventive measures will be shown as the final outcome.

4) The External actor is given a role in architecture so that they can monitor the patient’s health growth chart.

III. ALGORITHM

A. Decision Tree

A Decision tree is a type of supervised learning algorithm that is mostly used in classification problems. It represents a procedure for classifying categorical data based on their attributes. It is also efficient for the processing of a large amount of data, therefore, it is often used in data mining applications. The basic objective of this algorithm is to select the hierarchy of critical dependent variables that are relevant for the probabilities for particular business processes.

B. J48 Algorithm

The core goal of classification is to predict a category or class y from some inputs x. The applications of classification include sentiment analysis, ad targeting, spam detection, risk assessment, medical diagnosis and image classification. To generate decision trees, Ross Quinlan has developed an algorithm named as C4.5. It is an extension of Quinlan's earlier ID3 algorithm. The decision trees generated by C4.5 can be used for classification, and for this reason, C4.5 is often referred to as a statistical classifier. J48 is an open source Java implementation of the C4.5 algorithm [17]. The basic idea is to divide the data into range based on the attribute values for that item that are found in the training sample.

The main challenge here is to select an attribute for first split criteria & to read this attribute; we need to calculate information gain. Information gain represents the amount of information that would be needed to place a new instance in the branch. It is a difference between the entropy before and after a decision. The one with more information gain is considered as a valid attribute for first split criteria. In short, J48 has used information gain to decide the attribute that best fits for target variable prediction [20].

C. Counting gain

Entropy is a measure of the uncertainty associated with a random variable. This process uses the entropy which is a measure of the data disorder. The Entropy is calculated by,

$$\boldsymbol{Entropy}(\vec{y}) = -\sum_{j=1}^{n} \frac{|y_i|}{|\vec{y}|} \log\left(\frac{|y_i|}{|\vec{y}|}\right)$$

$$\boldsymbol{Entropy}(j|\vec{y}) = \frac{|y_j|}{|\vec{y}|} \log\left(\frac{|y_j|}{|\vec{y}|}\right)$$

As the data become purer and purer, the entropy value becomes smaller and smaller.

And the information gain is calculated by,

$$\boldsymbol{Gain}(\vec{y}, j) = \boldsymbol{Entropy}(\vec{y} - \boldsymbol{Entropy}(j|\vec{y}))\,|$$

The objective is to maximize the gain dividing by overall entropy due to a split argument by value j.

D. Pruning

Pruning is cutting away dead & overgrown branches. A bigger decision tree often puts us to the threat of overfitting model which results in a poor predictive performance. Overfitting often occurs when the model is overly complex & thereby increasing the chances of noise & error. The pruning is performed for decreasing classification errors which are being produced by specialization in the training set. Pruning can either be done in a top down or bottom up fashion.

There are two strategies of pruning, postpruning & prepruning. Postpruning is to take a fully grown tree & discard unreliable parts. Prepruning is to stop growing a branch when information becomes unreliable [14].

E. Limitations

The run-time complexity of this algorithm matches to the tree depth, which cannot be greater than the number of attributes. Tree depth is linked to tree size, and thereby to the number of examples. So, the size of C4.5 trees increases linearly with the number of examples. C4.5 rules slow for large and noisy datasets. Space complexity is very large as we have to store the values repeatedly in arrays. C4.5 rules slow for the large & noisy dataset [15].

F. Related Work

Here, we have used the J48 algorithm to predict the type of diabetes i.e. type1 or type2 based on the following data:

TABLE II. DIABETES DATA

Sr.No	Title	Description
1.	Attributes considered for prediction:	Age, Sex, FBSL, PPBSL, Height, Weight, HbA1c, Hb, TC, VLDL, Urine Protine, Urine Microalbumin, HDL, Urine Ketone, TG, Peripheral Pulses, Total Cholesterol, BMI, SBP & DBP
2.	Training data count:	2051 patients' data is processed.

1) Type1 diabetes: The body does not produce insulin. People usually develop type 1 diabetes before their 40th year, often in early adulthood or teenage years [18].

2) Type2 diabetes: The body does not produce enough insulin for proper function, or the cells in the body do not react to insulin [18].

The diagnosis of diabetes in patients is calculated by using the decision tree in two phases: data pre-processing in which the attributes are identified and second is diabetes prediction model constructed with the help of using the decision tree method. Both the phases are implemented using WEKA data mining tool. The performance comparison of decision tree algorithms and artificial neural network on medical data was performed on the basis of parameters as kappa statistics, mean absolute error, relative squared error, time to model and mean-squared error. On the basis of results, it has been examined that decision tree algorithm performs better than the artificial neural network. The overall accuracy is around 83 percent [19]. Decision tree method is chosen as the optimal for the problem as it has shown better results than other algorithms. J48, Random Forest, Naive Bayes etc. algorithms are used for disease diagnosis as they led to good accuracy. They were used to make predictions [19].

IV. REQUIREMENT, ANALYSIS & OUTPUT

A. Performance Requirement

Errors during the ordering or administration process account for about 90% of preventable adverse events. EHR (Electronic Health Record), defined as the process of entering medication orders electronically instead of on paper charts has been considered and promoted as an important tool to improve the quality and safety of drug prescription and thus reduce errors.

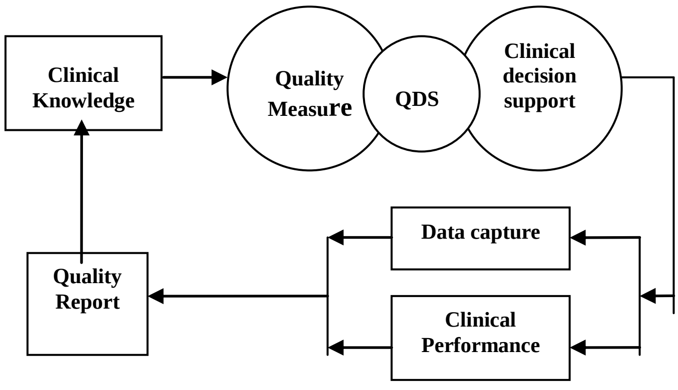

Fig.3. High-level overview of the relationship among CDSS, quality measures, and the QDS Model.

The Performance of any system is greatly impacted by the quality of its data set (QDS). Good Quality data can gain insight about forecasting decisions that derive accurate assistance at the time of care. As shown in Fig.3, a result is a quality report that indicates the level of performance against quality measures and provides additional knowledge. Data quality is an evaluation of appropriateness & relevance of data. Correctness, accuracy, completeness, timeliness & metadata are five measures of data quality that can be used for quality care in given framework.

B. System Quality Attributes

1) Decreasing errors of omission & commission.
2) Reducing unnecessary, ineffective, harmful care.
3) Adherence to evidence-based care.

Machine Learning is used when it is not feasible to solve complex problems by writing traditional algorithms. Problems like assisting physicians at the point of care can't be solved by the static or universal logic of algorithms but can be solved by specifying data-driven criterion for decision making. Mistakes can never be tolerated when it is a subject of health care & there is a possibility that misdiagnosis might happen through CDSS as well. To avoid this, we need to access system quality attributes by analyzing the differences between predicted & actual outcome.

C. Safety Requirement

Patient safety holds the topmost concern in CDSS & hence the requirements are given as follows [13]:

TABLE III. PATIENT SAFETY REQUIREMENT

Sr.No	Target Area of Care	Safety Requirements
1.	Preventive care	Immunization, screening, disease management guidelines for secondary prevention.
2.	Diagnosis	Suggestions for possible diagnoses that match a patient's signs and symptoms.
3.	Planning & implementation requirement	Treatment guidelines for specific diagnoses, drug dosage recommendations, alerts for drug-drug interactions.
4.	Followup management	Corollary orders, reminders for drug adverse event monitoring.
5.	Hospital provider efficiency	Care plans to minimize length of stay, order sets.
6.	Cost reduction & improved patient convenience	Duplicate testing alerts, drug formulary guidelines.

D. Decision Analysis

Clinical Decision analysis is a quantitative evaluation of the outcome that results from a set of choices in a specific clinical situation and is the methodology usually chosen for clinical problem-solving. This process is often implicit and occurs within internal algorithms and heuristics (mental shortcuts) that the clinician has developed and acquired over time. Decision analysis by requiring a specific model structure and assessment of various likelihoods and values of the outcomes makes the decision process explicit and more inclined to examination, discussion, and re-validation.

Decision models are often used as analytic tools to conduct cost-effectiveness analyses since this methodology can be used to find the expected value of predicted outcomes.

In general, decision analyses have been developed to:

1) Assist in clinical decision-making for a specific individual patient.

2) Estimate optimal strategies for classes of patients with specific clinical characteristics in given situations.

3) Link estimates of both clinical and economic outcomes (cost-effectiveness analysis) to help in forming issues regarding health policy.

4) Provide estimates of expected outcomes in situations.

E. Output

Following are the output snaps of a CDSS system developed by us. It has two separate modules, doctor & patient namely.

Fig.4. Add a new patient

Fig.4. shows the page for adding a new patient for the diagnosis. The doctor has to add the symptoms along with their values of a patient.

Fig.5. T1DM (Diagnosis)

Fig.5. shows the predicted disease (Type1-diabetes) based on the information of a patient added by the doctor. It also provides two options i.e. Get report and get treatment.

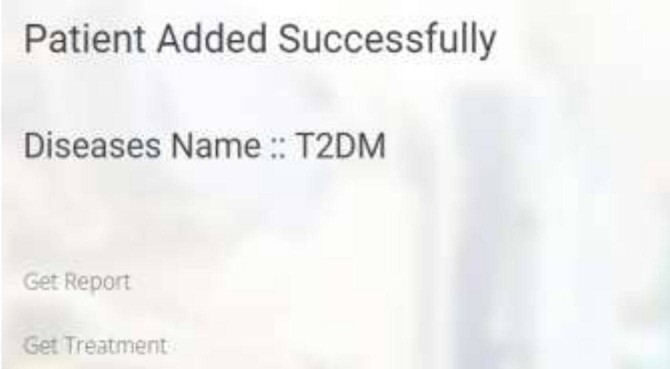

Fig.6. T2DM (Diagnosis)

Fig.6. shows the predicted disease (Type2-diabetes) based on the information of a patient added by the doctor.

Fig.7. Patient History

Fig.7. shows that the doctor can search for a particular patient. If the patient is already registered with this doctor then his id, name, disease, symptoms, medicine prescribed and the date when the doctor registered the patient is being displayed.

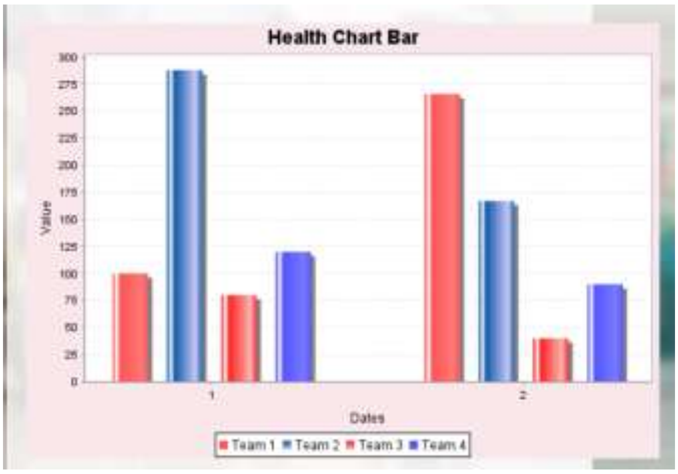

Fig.8. Patient Health Graph

Fig.8. shows that the patient or his relatives can log in to the portal with credentials given & can have a look at patients health chart.

V. CDSS FUTURE SCOPE AND APPLICATIONS

There are multiple clinical tasks that CDSS can accomplish or support in. Although the uncertainty about the business value of CDSS investment, the health care organizations will be in need of acquiring expertise technology for patient safety [16]. Future scope of CDSS is given as follows:

1) Image Realization and Interpretation: Critical diseases which require diagnosis on the basis of images such as x-rays, ultrasound, CT-Scan, MRI, etc. can also be predicted by using technologies like image processing, neural networks etc. The greatest benefit can be seen when examining a series of images where a system may detect minute changes over time that may have been overlooked by a professional due to the minuteness.

2) Pharmaceutical companies: They can use predictive analytics to best meet the needs of the public for medications.

VI. CONCLUSION

When it comes to the healthcare service delivery to the patients, it becomes important for all developing nations to pursue incremental improvement with respect to the faster diagnosis and treatment facilities. To subdue this need for innovation, CDSS will provide clinicians and practitioners with accurate assessments to improve clinical decision making, therefore, reducing serious medication errors. Ongoing efforts are geared towards increasing the size of data set as in to increase the number of critical diseases to be predicted. In this paper we presented the CDSS with analysis of one of the critical disease, working of an algorithm used for the prediction and state in research to expand the scope of the proposed system. This system results in improved efficiency in health care delivery e.g. by reducing costs through faster decision making, reductions in test duplication, decreased errors, involving patient's participation in reviewing the growth rate of their health and changed patterns of recommendation etc. The system is of great relevance to the user in providing the correct medication about him/her and improved healthcare facilities.

ACKNOWLEDGMENT

We take this opportunity to express our gratitude and deep regards to our Head of Department and our guide Prof. S.M Bhagat for his exemplary guidance and encouragement throughout this paper work. We extend our deep sense of appreciation to all those who had helped us to make this paper possible. Hereby, we truly appreciate the blessings, help, and guidance given by our professors & also the opportunity given by IERJ. The opportunity taken by us shall carry us to the long way in the journey of life on which we are about to embark.

REFERENCES

[1].ETA S. Berner, "CDSS: Theory & Practice", Science Business Media, LLC, ISBN-10:0-387-33914, 2007.

[2].Mohammed Ghazal, Yasmina Al Khalil, Fatemeh Jalil Dehbozorgi, and Marah Talal Alhalabi, "An Integrated Caregiver-focused mH health Framework for Elderly Care.", Journal of e-Health Pervasive Wireless Application & services, 2015.

[3].Dramane Sanou, "Ethical Issues and Safety in the Use of Clinical Decision Support Systems (CDSS)".

[4].Kensaku Kawamoto, Caitlin A Houlihan, E Andrew Balas, David F Lobach, "Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success".

[5].Bonnie Kaplan, "Evaluating informatics applications clinical decision support systems literature review", International Journal of Medical Informatics 64: 15-37, 5 July 2001.

[6].Eta S. Berner, "Clinical Decision Support Systems: State of the Art", AHRQ, Publication No.09-0069-EF, June 2009.

[7].Enrico Coiera, "Clinical Decision Support Systems".

[8].M.M.Abbasi, S. Kashiyarndi, "Clinical Decision Support Systems: A discussion on different methodologies used in Health Care".

[9].Alan A Montgomery, Tom Fahey, Tim J Peters, Christopher MacIntosh, Deborah J Sharp, "Evaluation of computer based clinical decision support system and risk chart for management of hypertension in primary care: randomized controlled trial".

[10].https://en.wikipedia.org/wiki/Clinical_decision_support_system

[11].http://timesofindia.indiatimes.com/india/India-records-5-2-million-medical-injuries-a-year/articleshow/22832260.cms

[12].http://www.euro.who.int/en/health-topics/Health-systems/patient-safety/data-and-statistics

[13].Prabhu Murugesan, Senthil Prabhu Natarajan, Lakshmi Karthigeyan, "Clinical decision support system", Computer Science Corporation, 2014.

[14].https://www.slideshare.net/marinasantini1/lecture-4-decision-trees-2-entropy-information-gain-gain-ratio-55241087

[15].Dr. Neeraj Bhargava, Girja Sharma, Dr. Ritu Bhargava, Manish Mathuria, "Decision Tree Analysis on J48 Algorithm for Data Mining", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 6, June 2013 ISSN: 2277 128X

[16].https://www.alexandria.unisg.ch/214544/1/ISHIMR2008.pdf

[17].https://en.wikipedia.org/wiki/C4.5_algorithm

[18].http://www.medicalnewstoday.com/info/diabetes

[19]. Gaganjot Kaur, Amit Chhabra, "Improved J48 Classification Algorithm for the Prediction of Diabetes", International Journal of Computer Applications (0975 – 8887) Volume 98 – No.22, July 2014.

[20].http://data-mining.business-intelligence.uoc.edu/ home/ j48-decision-tree

