ISSN (ONLINE) : 2045-8711
ISSN (PRINT) : 2045-869X

INTERNATIONAL JOURNAL OF INNOVATIVE TECHNOLOGY & CREATIVE ENGINEERING

March 2023
Vol - 13 No - 3
@IJITCE Publication

Design by Hemharsith

UK: Managing Editor
International Journal of Innovative Technology and Creative Engineering
1a park lane,
Cranford
London
TW59WA
UK

USA: Editor
International Journal of Innovative Technology and Creative Engineering
Dr. Arumugam
Department of Chemistry
University of Georgia
GA-30602, USA.

India: Editor
International Journal of Innovative Technology & Creative Engineering
36/4 12th Avenue,
1st cross St,
Vaigai Colony
Ashok Nagar
Chennai, India 600083

Email: editor@ijitce.co.uk

www.ijitce.co.uk

IJITCE PUBLICATION

International Journal of Innovative Technology & Creative Engineering

Vol.13 No.03

March 2023

www.ijitce.co.uk

Dear Researcher,

Greetings!

Articles in this issue discusses about study on data mining concepts in computer science.

We look forward many more new technologies in the next month.

Thanks,
Editorial Team
IJITCE

Editorial Members

Dr. Chee Kyun Ng Ph.D
Department of Computer and Communication Systems,
Faculty of Engineering,Universiti Putra Malaysia,UPMSerdang, 43400 Selangor,Malaysia.

Dr. Simon SEE Ph.D
Chief Technologist and Technical Director at Oracle Corporation, Associate Professor (Adjunct) at Nanyang Technological University
Professor (Adjunct) at ShangaiJiaotong University, 27 West Coast Rise #08-12,Singapore 127470

Dr. sc.agr. Horst Juergen SCHWARTZ Ph.D,
Humboldt-University of Berlin,Faculty of Agriculture and Horticulture,Asternplatz 2a, D-12203 Berlin,Germany

Dr. Marco L. BianchiniPh.D
Italian National Research Council; IBAF-CNR,Via Salaria km 29.300, 00015 MonterotondoScalo (RM),Italy

Dr. NijadKabbara Ph.D
Marine Research Centre / Remote Sensing Centre/ National Council for Scientific Research,
P. O. Box: 189 Jounieh,Lebanon

Dr. Aaron Solomon Ph.D
Department of Computer Science,
National Chi Nan University,No. 303, University Road,Puli Town, Nantou County 54561,Taiwan

Dr. Arthanariee. A. M M.Sc.,M.Phil.,M.S.,Ph.D
Director - Bharathidasan School of Computer Applications, Ellispettai, Erode, Tamil Nadu,India

Dr. Takaharu KAMEOKA, Ph.D
Professor, Laboratory of Food,
Environmental & Cultural Informatics Division of Sustainable Resource Sciences,
Graduate School of Bioresources,Mie University, 1577 Kurimamachiya-cho, Tsu, Mie, 514-8507, Japan

Dr. M. Sivakumar M.C.A.,ITIL.,PRINCE2.,ISTQB.,OCP.,ICP. Ph.D.
Technology Architect, Healthcare and Insurance Industry, Chicago, USA

Dr. Bulent AcmaPh.D
Anadolu University,
Department of Economics,Unit of Southeastern Anatolia Project(GAP),26470 Eskisehir,TURKEY

Dr. Selvanathan Arumugam Ph.D
Research Scientist, Department of Chemistry, University of Georgia, GA-30602,USA.

Dr. S.Prasath Ph.D
Assistant Professor, School of Computer Science, VET Institute of Arts & Science (Co-Edu) College, Erode, Tamil Nadu, India

Dr. P.Periyasamy, M.C.A.,M.Phil.,Ph.D.
Associate Professor, Department of Computer Science and Applications, SRM Trichy Arts and Science College, SRM Nagar, Trichy - Chennai Highway, Near Samayapuram, Trichy - 621 105,

Mr. V N Prem Anand
Secretary, Cyber Society of India

Review Board Members

Dr. Rajaram Venkataraman
Chief Executive Officer, Vel Tech TBI || Convener, FICCI TN State Technology Panel || Founder, Navya Insights || President, SPIN Chennai

Dr. Paul Koltun
Senior Research ScientistLCA and Industrial Ecology Group,Metallic& Ceramic Materials,CSIRO Process Science & Engineering Private Bag 33, Clayton South MDC 3169,Gate 5 Normanby Rd., Clayton Vic. 3168, Australia

Dr. Zhiming Yang MD., Ph. D.
Department of Radiation Oncology and Molecular Radiation Science,1550 Orleans Street Rm 441, Baltimore MD, 21231,USA

Dr. Jifeng Wang
Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign Urbana, Illinois, 61801, USA

Dr. Giuseppe Baldacchini
ENEA - Frascati Research Center, Via Enrico Fermi 45 - P.O. Box 65,00044 Frascati, Roma, ITALY.

Dr. MutamedTurkiNayefKhatib
Assistant Professor of Telecommunication Engineering,Head of Telecommunication Engineering Department,Palestine Technical University (Kadoorie), TulKarm, PALESTINE.

Dr.P.UmaMaheswari
Prof &Head,Depaartment of CSE/IT, INFO Institute of Engineering,Coimbatore.

Dr. T. Christopher, Ph.D.,
Assistant Professor &Head,Department of Computer Science,Government Arts College(Autonomous),Udumalpet, India.

Dr. T. DEVI Ph.D. Engg. (Warwick, UK),
Head,Department of Computer Applications,Bharathiar University,Coimbatore-641 046, India.

Dr. Renato J. orsato
Professor at FGV-EAESP,Getulio Vargas Foundation,São Paulo Business School,RuaItapeva, 474 (8° andar),01332-000, São Paulo (SP), Brazil
Visiting Scholar at INSEAD,INSEAD Social Innovation Centre,Boulevard de Constance,77305 Fontainebleau - France

Y. BenalYurtlu
Assist. Prof. OndokuzMayis University

Dr.Sumeer Gul
Assistant Professor,Department of Library and Information Science,University of Kashmir,India

Dr. ChutimaBoonthum-Denecke, Ph.D
Department of Computer Science,Science& Technology Bldg., Rm 120,Hampton University,Hampton, VA 23688

Dr. Renato J. Orsato
Professor at FGV-EAESP,Getulio Vargas Foundation,São Paulo Business SchoolRuaItapeva, 474 (8° andar),01332-000, São Paulo (SP), Brazil

Dr. Lucy M. Brown, Ph.D.
Texas State University,601 University Drive,School of Journalism and Mass Communication,OM330B,San Marcos, TX 78666

JavadRobati
Crop Production Departement,University of Maragheh,Golshahr,Maragheh,Iran

VineshSukumar (PhD, MBA)
Product Engineering Segment Manager, Imaging Products, Aptina Imaging Inc.

Dr. Binod Kumar PhD(CS), M.Phil.(CS), MIAENG,MIEEE
Professor, JSPM's Rajarshi Shahu College of Engineering, MCA Dept., Pune, India.

Dr. S. B. Warkad
Associate Professor, Department of Electrical Engineering, Priyadarshini College of Engineering, Nagpur, India

Dr. doc. Ing. RostislavChoteborský, Ph.D.
Katedramateriálu a strojírenskétechnologie Technickáfakulta,Ceskázemedelskáuniverzita v Praze,Kamýcká 129, Praha 6, 165 21

Dr. Paul Koltun
Senior Research ScientistLCA and Industrial Ecology Group,Metallic& Ceramic Materials,CSIRO Process Science & Engineering Private Bag 33, Clayton South MDC 3169,Gate 5 Normanby Rd., Clayton Vic. 3168

DR.ChutimaBoonthum-Denecke, Ph.D
Department of Computer Science,Science& Technology Bldg.,HamptonUniversity,Hampton, VA 23688

Mr. Abhishek Taneja B.sc(Electronics),M.B.E,M.C.A.,M.Phil.,
Assistant Professor in the Department of Computer Science & Applications, at Dronacharya Institute of Management and Technology, Kurukshetra. (India).

Dr. Ing. RostislavChotěborský,ph.d,
Katedramateriálu a strojírenskétechnologie, Technickáfakulta,Českázemědělskáuniverzita v Praze,Kamýcká 129, Praha 6, 165 21

Dr. AmalaVijayaSelvi Rajan, B.sc,Ph.d,
Faculty – Information Technology Dubai Women's College – Higher Colleges of Technology,P.O. Box – 16062, Dubai, UAE

Naik Nitin AshokraoB.sc,M.Sc
Lecturer in YeshwantMahavidyalayaNanded University

Dr.A.Kathirvell, B.E, M.E, Ph.D,MISTE, MIACSIT, MENGG
Professor - Department of Computer Science and Engineering,Tagore Engineering College, Chennai

Dr. H. S. Fadewar B.sc,M.sc,M.Phil.,ph.d,PGDBM,B.Ed.
Associate Professor - Sinhgad Institute of Management & Computer Application, Mumbai-BangloreWesternly Express Way Narhe, Pune - 41

Dr. David Batten
Leader, Algal Pre-Feasibility Study,Transport Technologies and Sustainable Fuels,CSIRO Energy Transformed Flagship Private Bag 1,Aspendale, Vic. 3195,AUSTRALIA

Dr R C Panda
(MTech& PhD(IITM);Ex-Faculty (Curtin Univ Tech, Perth, Australia))Scientist CLRI (CSIR), Adyar, Chennai - 600 020,India

Miss Jing He
PH.D. Candidate of Georgia State University,1450 Willow Lake Dr. NE,Atlanta, GA, 30329

Jeremiah Neubert
Assistant Professor,MechanicalEngineering,University of North Dakota

Hui Shen
Mechanical Engineering Dept,Ohio Northern Univ.

Dr. Xiangfa Wu, Ph.D.
Assistant Professor / Mechanical Engineering,NORTH DAKOTA STATE UNIVERSITY

SeraphinChallyAbou
Professor,Mechanical& Industrial Engineering Depart,MEHS Program, 235 Voss-Kovach Hall,1305 OrdeanCourt,Duluth, Minnesota 55812-3042

Dr. Qiang Cheng, Ph.D.
Assistant Professor,Computer Science Department Southern Illinois University CarbondaleFaner Hall, Room 2140-Mail Code 45111000 Faner Drive, Carbondale, IL 62901

Dr. Carlos Barrios, PhD
Assistant Professor of Architecture,School of Architecture and Planning,The Catholic University of America

Y. BenalYurtlu
Assist. Prof. OndokuzMayis University

Dr. Lucy M. Brown, Ph.D.
Texas State University,601 University Drive,School of Journalism and Mass Communication,OM330B,San Marcos, TX 78666

Dr. Paul Koltun
Senior Research ScientistLCA and Industrial Ecology Group,Metallic& Ceramic Materials CSIRO Process Science & Engineering

Dr.Sumeer Gul
Assistant Professor,Department of Library and Information Science,University of Kashmir,India

Dr. ChutimaBoonthum-Denecke, Ph.D
Department of Computer Science,Science& Technology Bldg., Rm 120,Hampton University,Hampton, VA 23688

Dr. Renato J. Orsato
Professor at FGV-EAESP,Getulio Vargas Foundation,São Paulo Business School,RuaItapeva, 474 (8° andar)01332-000, São Paulo (SP), Brazil

Dr. Wael M. G. Ibrahim
Department Head-Electronics Engineering Technology Dept.School of Engineering Technology ECPI College of Technology 5501 Greenwich Road - Suite 100,Virginia Beach, VA 23462

Dr. Messaoud Jake Bahoura
Associate Professor-Engineering Department and Center for Materials Research Norfolk State University,700 Park avenue,Norfolk, VA 23504

Dr. V. P. Eswaramurthy M.C.A., M.Phil., Ph.D.,
Assistant Professor of Computer Science, Government Arts College(Autonomous), Salem-636 007, India.

Dr. P. Kamakkannan,M.C.A., Ph.D .,
Assistant Professor of Computer Science, Government Arts College(Autonomous), Salem-636 007, India.

Dr. V. Karthikeyani Ph.D.,
Assistant Professor of Computer Science, Government Arts College(Autonomous), Salem-636 008, India.

Dr. K. Thangadurai Ph.D.,
Assistant Professor, Department of Computer Science, Government Arts College (Autonomous), Karur - 639 005,India.

Dr. N. Maheswari Ph.D.,
Assistant Professor, Department of MCA, Faculty of Engineering and Technology, SRM University, Kattangulathur, Kanchipiram Dt - 603 203, India.

Mr. Md. Musfique Anwar B.Sc(Engg.)
Lecturer, Computer Science & Engineering Department, Jahangirnagar University, Savar, Dhaka, Bangladesh.

Mrs. Smitha Ramachandran M.Sc(CS).,
SAP Analyst, Akzonobel, Slough, United Kingdom.

Dr. V. Vallimayil Ph.D.,
Director, Department of MCA, Vivekanandha Business School For Women, Elayampalayam, Tiruchengode - 637 205, India.

Mr. M. Moorthi M.C.A., M.Phil.,
Assistant Professor, Department of computer Applications, Kongu Arts and Science College, India

PremaSelvarajBsc,M.C.A,M.Phil
Assistant Professor,Department of Computer Science,KSR College of Arts and Science, Tiruchengode

Mr. G. Rajendran M.C.A., M.Phil., N.E.T., PGDBM., PGDBF.,
Assistant Professor, Department of Computer Science, Government Arts College, Salem, India.

Dr. Pradeep H Pendse B.E.,M.M.S.,Ph.d
Dean - IT,Welingkar Institute of Management Development and Research, Mumbai, India

Muhammad Javed
Centre for Next Generation Localisation, School of Computing, Dublin City University, Dublin 9, Ireland

Dr. G. GOBI
Assistant Professor-Department of Physics,Government Arts College,Salem - 636 007

Dr.S.Senthilkumar
Post Doctoral Research Fellow, (Mathematics and Computer Science & Applications),UniversitiSainsMalaysia,School of Mathematical Sciences, Pulau Pinang-11800,[PENANG],MALAYSIA.

Manoj Sharma
Associate Professor Deptt. of ECE, PrannathParnami Institute of Management & Technology, Hissar, Haryana, India

RAMKUMAR JAGANATHAN
Asst-Professor,Dept of Computer Science, V.L.B Janakiammal college of Arts & Science, Coimbatore,Tamilnadu, India

Dr. S. B. Warkad
Assoc. Professor, Priyadarshini College of Engineering, Nagpur, Maharashtra State, India

Dr. Saurabh Pal
Associate Professor, UNS Institute of Engg. & Tech., VBS Purvanchal University, Jaunpur, India

Manimala
Assistant Professor, Department of Applied Electronics and Instrumentation, St Joseph's College of Engineering & Technology, Choondacherry Post, Kottayam Dt. Kerala -686579

Dr. Qazi S. M. Zia-ul-Haque
Control Engineer Synchrotron-light for Experimental Sciences and Applications in the Middle East (SESAME),P. O. Box 7, Allan 19252, Jordan

Dr. A. Subramani, M.C.A.,M.Phil.,Ph.D.
Professor,Department of Computer Applications, K.S.R. College of Engineering, Tiruchengode - 637215

Dr. SeraphinChallyAbou
Professor, Mechanical & Industrial Engineering Depart. MEHS Program, 235 Voss-Kovach Hall, 1305 Ordean Court Duluth, Minnesota 55812-3042

Dr. K. Kousalya
Professor, Department of CSE,Kongu Engineering College,Perundurai-638 052

Dr. (Mrs.) R. Uma Rani
Asso.Prof., Department of Computer Science, Sri Sarada College For Women, Salem-16, Tamil Nadu, India.

MOHAMMAD YAZDANI-ASRAMI
Electrical and Computer Engineering Department, Babol"Noshirvani" University of Technology, Iran.

Dr. Kulasekharan, N, Ph.D
Technical Lead - CFD,GE Appliances and Lighting,
GE India,John F Welch Technology Center,Plot # 122, EPIP, Phase 2,Whitefield Road,Bangalore – 560066, India.

Dr. Manjeet Bansal
Dean (Post Graduate),Department of Civil Engineering,Punjab Technical University,GianiZail Singh Campus,Bathinda -151001 (Punjab),INDIA

Dr. Oliver Jukić
Vice Dean for education,Virovitica College,MatijeGupca 78,33000 Virovitica, Croatia

Dr. Lori A. Wolff, Ph.D., J.D.
Professor of Leadership and Counselor Education,The University of Mississippi,Department of Leadership and Counselor Education, 139 Guyton University, MS 38677

Contents

A SURVEY ON VARIOUS OSS TOOLS AND TECHNIQUES
IN RECENT MEDICAL FIELD OF DATA MINING ……..…..…. [1178]

A SURVEY ON VARIOUS OSS TOOLS AND TECHNIQUES IN RECENT MEDICAL FIELD OF DATA MINING

M.Baskar[1]
Ph.D Research Scholar, Dept. of Computer Science and Applications,
Park College of Arts and Science, Chinnakarai Tirupur, Tamilnadu,India
baskarm2cs@gmail.com.
Dr.P.Periyasamy[2]
Associate professor of Computer Science and Applications,
Park College of Arts and Science, Chinnakarai Tirupur, Tamilnadu,India
periyasamy_p_2007@yahoo.co.in.

***Abstract* – Data mining is the one of the hottest topics in research of computer Science. In medical field the industry gathers massive amount of healthcare data which are not "mined" to find out unknown information. Data Mining tools and techniques can be positively applied in many fields in various kinds of . Now a days so many people are also suffering various kinds of diseases. The medicinal industries come across with new treatments and medicine each and every day. The medicinal industries must deliver well conclusion and remedy to the patients to achieve worthy quality of service. Most of the Organizations using Data Mining as a powerful tool, to deal with the reasonable situation for data analysis. The primary purpose of this paper is to provide detail information about which tools and techniques are used to identify the accuracy level of various diseases and data mining applications. This data mining based prediction system reduces the human effects and cost effective one. There are various kinds of oss tools mining tools are discussed.**

Keywords - Data Mining: concepts-Tools and Techniques-Algorithms Health Data Analysis-Data Mining Applications; Classification; Clustering; Association.

1. INTRODUCTION

Data mining is also known as knowledge discovery from Data (KDD). The purpose of data mining is to mine useful information from huge databases or data warehouses. Now a days, Data Mining is becoming common in healthcare field because there is an essential of operational analytical methodology for detecting unidentified and valuable information in health data. In health industry, Data Mining offers numerous benefits such as recognition of the fraud in health insurance, availability of medical solution to the patients at lesser price, detection of bases of diseases and identification of medical treatment methods. Data mining algorithms useful in healthcare industry and shows an important role in estimate and finding of the diseases[1]. There are a huge number of data mining applications are establish in the medical related areas such as Medical device industry, Pharmaceutical Industry and Hospital Management and health sector management. To catch the valuable and unknown information from the database is the determination behind the application of data mining. The knowledge discovery is an interactive process, containing by developing an understanding of the application domain, choosing and making a data set, preprocessing, data transformation[2]

The data made by the health organizations is exact huge and difficult due to which it is hard to analyze the data in order to mark vital conclusion regarding patient health. This data covers details regarding hospitals, patients, medical claims, treatment cost and etc. So, there is an essential to make a powerful tool for analyzing and extracting significant information from this complex data. The analysis of health data expands the healthcare by improving the presentation of patient management jobs[3]. The consequence of Data Mining technologies are to make available welfares to healthcare organization for grouping the patients having related/similar type of diseases or health issues so that healthcare organization provides them active treatments. It can also valuable for predicting the how many days of stay of patients in hospital, for medical diagnosis and creating plan

for active information system management. New and current technologies are used in medical field to improve the medical services in cost effective manner[4].

2. Classification of Data Mining System

Data mining systems can be categorized according to various criteria as given below

1) Type of data sources mined
2) Database involved
3) Kind of knowledge discovered
4) Mining techniques used

3. Process of Data Mining

The Data mining process includes the following few steps

1) Data Cleaning: It is used to remove noise and inconsistent data.
2) Data Integration: It is used to combine multiple data sources.
3) Data Selection: It is used to retrieve the relevant data from the database for analysis task.
4) Data Transformation: It is used to transformed or consolidated data into particular appropriate form for mining by performing summary or aggregation operations.
5) Data Mining: Here the intelligent methods are applied in order to extract data patterns.
6) Pattern Evaluation: It is used to evaluate the data patterns.
7) Knowledge Presentation: Here the knowledge is represented.
5) Data Mining: Here the intelligent methods are applied in order to extract data patterns.
6) Pattern Evaluation: It is used to evaluate the data patterns.
7) Knowledge Presentation: Here the knowledge is represented.

The following figure 1 shows the data mining process.

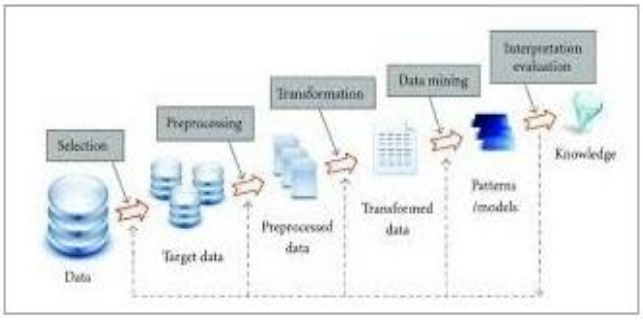

Figure: 1 Data Mining steps

3.1. Data mining techniques

There are enormous number of data mining techniques have been evolving and using in data mining projects recently. Some of the data mining techniques are given below,

3.1.1 Association

Association is one of the top - well known data mining techniques. In association, a pattern is learned based on an association between items in the similar transaction. That's the purpose the association technique is also well-known as relation technique. The association technique is used in marketplace basket analysis to classify a set of products that customers regularly purchase together[5]. Dealers are using association technique to investigation buyer's buying lifestyles. Based on ancient sale data, retailers might catch out that customers always buy jam when they buy breads, and, therefore, they can put jam and breads following to each other to save time for customer and make steps to growth sale.

3.1.2 Classification

Classification is a common data mining technique based on machine learning. Mostly, classification is used to categorize each item in a set of data into one of a predefined set of classes or groups[6].

Classification method uses variety of mathematical techniques such as decision trees, linear programming, neural network and statistics. In classification, make the software that can acquire how to classify the data items into groups[6]. For instance, first apply classification in the application that "given all records of employees who left from the company; predict who will probably leave the company in a future period." In this case, we divide the records of employees into two groups that named "leave" and "stay". And then can ask our data mining software to classify the employees into separate groups[7].

3.1.3 Clustering

Clustering is a data mining technique that makes an expressive or valuable cluster of objects which have related or same characteristics using the automatic technique. The clustering technique describes the classes and puts objects in each class, while in the classification techniques,

objects are allotted into predefined classes. To make the concept richer, can take book management in the library as an instance. In a library, there is a wide variety of books on different topics available. The challenge is how to hold those books in a way that readers can take several books on a particular topic without trouble. By using the clustering technique, can retain books that have some kinds of similarities in one cluster or one bookshelf and label it with a meaningful name. If readers need to take books in that topic, they would only have to go to that bookshelf instead of looking for the whole library.

There are two types of cluster in data mining

i. Inter cluster
ii. Intra cluster

We have two type of distance – Intercluster Distance and Intracluster Distance**.**

Intercuster Distance:

Intercluster distance is the distance between two objects belonging to two different clusters.

Intracuster Distance:

Intracluster distance is the distance between two objects belonging to same cluster.

3.1.4 Prediction

The prediction, as its name implied, is one of a data mining techniques that determine the association between independent variables and correlation between dependent and independent variables. For instance, the prediction analysis technique can be used in the sale to calculate income for the future if consider the sale is an independent variable, income could be a dependent variable. Then based on the ancient sale and earnings data, we can draw a fixed regression curve that is used for profit prediction[8].

3.1.5 Sequential Patterns

Sequential patterns analysis is one of data mining technique that pursues to determine or recognize associated patterns, fixed events or fashions in transaction data over a business period. In sales, with ancient transaction data, businesses can recognize a set of items that customers buy together different times in a year. Then industries can use this information to mention customers buy it with better deals based on their purchasing regularity in the past[9].

3.1.6 Decision trees

One of the best classification techniques in data mining . Just like a hierarchical structure format. A decision tree is one of the best common used data mining techniques because its model is easy to understand for users. In decision tree technique, the root of the decision tree is a simple question or condition that has compound answers. Each answer then leads to a group of questions or conditions that helps to determine the data so that can make the ultimate decision based on it. For example, use the following decision tree to determine whether a person has eligible for vote or not[10].

Combine two or more of those data mining techniques composed to form an appropriate process that meets the business requirements.

Decision trees have three main parts: a root node, leaf nodes and branches. The root node is the starting point of the tree, and both root and leaf nodes contain questions or criteria to be answered. Branches are arrows connecting nodes, showing the flow from question to answer. Each node typically has two or more nodes extending from it. For example, if the question in the first node requires a "yes" or "no" answer, there will be one leaf node for a "yes" response, and another node for "no.

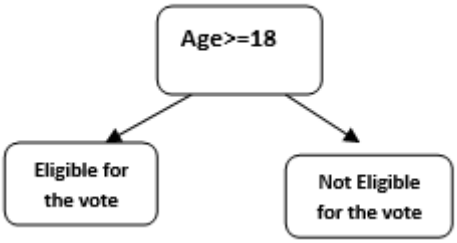

Fig.2 Decision Tree

3.1.8 Decision Tree uses

A decision tree can be used in either a predictive manner or a descriptive manner. In either instance they are constructed the same way and are always used to visualize all possible outcomes and decision points that occur chronologically[12]. Decision trees are most commonly used in the financial world for areas such as loan approval, portfolio management, and spending. A decision tree can also be helpful when examining the viability of a new product or defining a new market for an existing product[13].

4. Data Mining Tools

The Various kinds of data mining tools are given below for reference,

1) Artificial Neural Networks (ANN),
2) Rough Set Theory (RST),
3) Statistical Package for the Social Sciences modeler (SPSS),
4) K-means clustering
5) Single Nucleotide Polymorphism (SNP)
6) Orange
7) SAS Mining tools
8) Rattle
9) Rapid miner.
10) Data melt Data mining.

The Eight best open source Data mining tools are given below,

1) Tera data
2) R
3) Rapid miner
4) Weka
5) R-Programming 4) Orange
6) Knime
7) Natural Language Toolkit (NLTK)
8)Apache Mahout

R

R analytics is data analytics using R programming language, an open-source language used for statistical computing or graphics. This programming language is often used in statistical analysis and data mining. It can be used for analytics to identify patterns and build practical model

Rapid Miner

It is one of the best predictive analysis systems. Also, it was developed by the company with the same name as the Rapid Miner. It is written in JAVA programming language**.** It provides an integrated environment for deep learning.

Weka

This software developed at the University of Waikato in New Zealand. It is best suited for data analysis and predictive modeling. It contains algorithms and visualization tools that support machine learning**.**

Weka has a GUI that facilitates easy access to all its features. It is written in JAVA programming language**.**

KNIME

KNIME is the best integration platform for data analytics. Also reporting developed by KNIME.com AG. It operates on the concept of the modular data pipeline. KNIME constitutes of various machine learning and data mining components embedded together.

It has been used for pharmaceutical research. In addition, it performs for customer data analysis, financial data analysis.

Sisense

Sisense is extremely useful and best suited BI software. That it comes to reporting purposes within the organization. It is developed by the company of same name 'Sisense'. It has a brilliant capability to handle. Also, process data for the small-scale/large scale organizations.

SSDT

SSDT is a universal, declarative model. We use this model to expands all the phases of database development in the Visual Studio IDE. And developed to do data analysis and provide business intelligence solutions. Developers use SSDT transact- a design capability of SQL and refactor databases

SSDT Apache Mahout is a project developed by Apache Foundation. Also, it serves the primary purpose of creating machine learning algorithms**.** It focuses mainly on data clustering, classification, and collaborative filtering.

Mahout is written in JAVA and includes JAVA libraries to perform mathematical operations. Such as linear algebra and statistics. Mahout is growing continuously as the algorithms implemented inside Apache Mahout. The algorithms of Mahout have implemented a level above Hadoop.

Oracle

A component of Oracle Advanced Analytics, it software provides excellent data mining algorithms.

The algorithms designed inside ODM leverage the potential strengths of Oracle database. The data mining feature of SQL can dig data out of database tables, views, and schemas.

Rattle

A rattle is a GUI tool that uses R stats programming language**.** Rattle exposes the statistical power of R by providing considerable data mining functionality. Although Rattle has an extensive and well-developed UI. Also, it has an

inbuilt log code tab that generates duplicate code for any activity happening at GUI.

DataMelt, also known as DMelt is a computation and visualization environment. Also, provides an interactive framework to do data analysis and visualization. It is designed mainly for engineers, scientists & students.

DMelt is a multi-platform utility. It can run on any operating system which is compatible with JVM(Java Virtual Machine).

IBM Cognos BI is an intelligence suite. It consists of sub-components that meet specific organizational requirements.

Cognos Connection: A web portal to gather and summarize data in scoreboard/reports.

Orange

Orange is a perfect software suite for machine learning and data mining. It best aids the data visualization and is a component-based software.

As it is a software, the components of orange are called 'widgets'.

Widgets offer major functionalities like

- Showing data table and allowing to select features
- Reading the data
- Training predictors and to compare learning algorithms
- Visualizing data elements etc.

5. Data Mining Applications in Healthcare

Data mining tools are used to predict the effective results from the data verified on healthcare problems. Different data mining tools are used to calculate the accuracy level in different healthcare problems. The following list of medical problems has been evaluated and estimated[14].

- ✓ Heart Disease Cancer
- ✓ HIV/AIDS Blood
- ✓ Brain Cancer Tuberculosis
- ✓ Diabetes Mellitus Kidney dialysis
- ✓ Dengue IVF
- ✓ Hepatitis C

S.NO	TYPE OF DISEASES	DATA MINING TOOLS	TECHNIQUES	ALGORITHMS	Accuracy
1	Heart Diseases	H_2O	Classification	Navie	72.5
2	Cancer	Weka	Classification	Decision Table	95.0
3	Kidney Diseases	R	c.4.5	Decision making	89.4
4	Heart Diseases	H2O	Classification	Navie	80
5	Dengue	SPSS	Classification	C 5.0	91
6	Brain Cancer	K-Means	Clustering	j.48	73.20
7	Hepatitis C	Knime	Information Gain	Decision Rule	85

Table 1: Tools and Techniques in Healthcare

The above table 1 show the various tools and techniques are used to find the accuracy level of various diseases.

6. Conclusion

In this research paper mainly focus various OSS Software tools and Mining algorithms . The prediction of diseases using Data Mining applications is an inspiring task but it extremely reduces the human strength and increases the diagnostic accuracy. Developing well-organized data mining tools for an application could decrease the cost and time control in terms of human resources and capability. Discovering knowledge from the medical data is such a complicated job as the data found are noisy, irrelevant and massive too. In this scenario, data mining tools come in close in discovering of knowledge of the medical data and it is fairly interesting. It is detected from this table1 that a mixture of more than one data mining techniques than a single technique for diagnosing or predicting diseases in healthcare sector could produce additional encouraging results. The Table 1 displays the motivating results that data mining techniques in all the health care applications offer a more promising level of accuracy like 95.77% for cancer prediction and around 80 % for estimating the success rate of IVF treatment.

References

[1] M. Durai raj, and V. Ranjani, “Data Mining Applications in Healthcare Sector: A Study”, international journal of scientific & technology research volume 2, issue 10, October 2013 ISSN 2277- 8616

[2] HianChyeKoh and Gerald Tan, “Data Mining Applications in Healthcare journal of Healthcare Information Management” – Vol 19, No 2.

[3] JayanthiRanjan, “Applications of data mining techniques in pharmaceutical industry”, Journal of Theoretical and Applied Technology, (2007).

[4] RubanD.Canlas Jr., MSIT, MBA, “Data mining in Healthcare: Current applications and issues”.

[5] K. Srinivas, B. Kavitha Rani and Dr. A. Govidhan, “Applications of Data Mining Techniques in Healthcare and Prediction of Heart Attacks” International Journal on Computer Science and Engineering (2010).

[6] D.Shobana, N.Uthra,” The Data Mining Concepts and Techniques: A Survey” International Journal of Trend in Research and Development, Volume 2(6),ISSN 2394-9333

[7] Hwang, H.G., Chang, I.C., Xingquan Zhu, Ian Davidson, “Knowledge Discovery and Data Mining: Challenges and Realities”, ISBN 978-1-59904-252,Hershey, New York, 2007

[8] Han, J., Kamber, M., 2001.” Data Mining: Concepts and Techniques” , Morgan Kaufmann, San Fco., CA., USA [3].Witten, I. H., Frank, E. “Data mining: Practical machine learning tools and techniques with Javaimplementations”. Morgan Kaufmann, San Francisco, CA. USA, 371 pp, 2000

[9] K.Deepa “Rapid Development of Applications in Data Mining” Proceedings of 2013 International Conference on Green High Performance Computing March 14-15, 2013, India

[10] Introduction to Data Mining and Knowledge Discovery, Third Edition ISBN: 1-892095-02-5, Two Crows Corporation, 10500 Falls Road, Potomac, MD 20854 (U.S.A.), 1999.

[11] Han, J. & M. Kamber, and Data mining: concepts and techniques, San Francisco: Morgan Kaufman.

[12] Mahmoud H, Abbas E, Fathy I 2018 Data mining and ontology-based techniques in healthcare management *International Journal of Intelligent* Naveenkumar S, Kirubhakaran R, Jeeva G, Shobana M, Sangeetha K Smart Health Prediction Using Machine Learning

[13] Kumar H and Singh N 2017 Review paper on Big Data in healthcare informatics *International Research* Journal of Engineering and Technology 4(2) pp 197-201

[14] Balakrishnan S and Narayanaswamy R 2009 Feature selection using fcbf in type ii diabetes databases International Journal of the Computer the Internet and the Management 17(1) pp 50-8.

[15] Bahrami B and Shirvani M H 2015 Prediction and diagnosis of heart disease by data mining techniques Journal of Multidisciplinary Engineering Science and Technology (JMEST) 2(2) pp 164-8 Engineering Inforamtics.