

---

## A Review Based Study on Data Mining for Effective Decision Making in Rural Healthcare

Narendra Kumar Sharma\*

Dr. Shahnaz Fatima\*\*

Dr. Swati Saxena\*\*\*

---

### Abstract

The primary aim of this research paper is to study the literatures and their findings on data mining for effective decision making in health related issues specially for rural people. In our critical review, we compare verity of research papers based on data mining approaches, techniques, tools and its impact on healthcare. The research paper has alienated in five sections Introduction, Review of Literature, Rural Healthcare, Comparative Table and Conclusion. The researcher has tried to put on view all the existing findings to show the position of healthcare.

---

### Keywords:

Rural Healthcare;  
Decision Making;  
Medical Decision Support;  
Data Mining.

Copyright © 2017 International Journals of Multidisciplinary Research Academy. All rights reserved.

---

### Author correspondence:

Narendra Kumar Sharma  
Research Scholar, Amity University, Lucknow, Uttar Pradesh - India

---

### 1. Introduction

The People of India living in rural areas face major crisis due to shortage of proper healthcare institutions as well as medical professionals. The Municipal healthcare centre (MHC) and the public healthcare centre (PHC) are poorly run with less professional human resources and limited medical facilities. The rural medical health centre may have medications to provide first aid, but may not have the expertise. People of India in rural areas have to travel long distances to access proper health care which can be the difference between life and death in the case of emergency. There is lack of ambulance in the remote areas, even in urban areas to pinpoint the exact location of the patient is difficult.

### 2. Objective of Study:

To explore the existing research on data mining regarding healthcare system for future directions to others.

### 3. Rural Healthcare:

Healthcare is the right of every individual but lack of quality infrastructure, dearth of qualified medical functionaries and non- access to basic medicines and medical facilities thwarts its reach to 60% of population in India. A majority of 700 million people lives in rural areas where the condition of medical facilities is deplorable. Considering the picture of grim facts there is a dire need of new practices and procedures to ensure that quality and timely healthcare reaches the deprived corners of the Indian villages. Though a lot of

---

\*Narendra Kumar Sharma, Research Scholar, Amity University, Lucknow, Uttar Pradesh – India

\*\*Dr. Shahnaz Fatima, Assistant Professor III, Amity University, Lucknow, Uttar Pradesh – India

\*\*\*Dr. Swati Saxena, Assistant Professor, MPEC, Kanpur, Uttar Pradesh – India

policies and programs are being run by the Government but the success and effectiveness of these programs is questionable due to gaps in the implementation [11].

In rural India, where the number of Primary health care centers (PHCs) is limited, 8% of the centers do not have doctors or medical staff, 39% do not have lab technicians and 18% PHCs do not even have a pharmacist. India also accounts for the largest number of maternity deaths. A majority of these are in rural areas where healthcare situation is poor. Even in private sector, healthcare is often confined to family planning and antenatal care and do not extend to more critical services like labor and delivery, where proper medical care can save life in the case of complications [11].

#### 4. Review of Literature:

**4.1 M. Durairaj, V. Ranjani (2013)**, focused to compare a variety of techniques, approaches and different tools and its impact on the healthcare sector. they aims to make a detailed study report of different types of data mining applications in the healthcare sector and to reduce the complexity of the study of the healthcare data transactions and also presents a comparative study of different data mining applications, techniques and different methodologies applied for extracting knowledge from database generated in the healthcare industry [1].

**4.2 Boris Milovic and Milan Milovic (2012)**, discusses about information technologies and it's automatization of processes for extraction of data that help to get interesting knowledge and regularities, which means the elimination of manual tasks and easier extraction of data directly from electronic records, transferring onto secure electronic system of medical records which will save lives and reduce the cost of the healthcare services, as well and early discovery of contagious diseases with the advanced collection of data [2].

**4.3 Divya Tomar and Sonali Agarwal (2013)**, conduct a survey and explore the utility of various Data Mining techniques such as classification, clustering, association, regression in health domain and provide brief introduction of these techniques and their advantages and disadvantages. They also highlight applications, challenges and future issues of Data Mining in healthcare [3].

**4.4 Vishnu S. Pendyala, Yi Fang, JoAnne Holliday and Ali Zalzal (2014)**, presents our vision of automating some of the healthcare functions such as monitoring and diagnosis for mass deployment. They also explain our ideas on how machines can help in this essential life supporting activity [4].

**4.5 Shivani Batra, Shelly Sachdeva, Hem Jyotsana Parashar and Pulkit Mehndiratta (2013)**, compared the benefits gained from applying data mining techniques to standardized and non-standardized EHRs is provided. They aim to indentify the need of applying data mining techniques to standardized electronic healthcare records. It interrogates various issues that need to be resolved for providing an efficient standardized decision support system [5].

**4.6 Hian Chye Koh and Gerald Tan (2015)**, discusses data mining and its applications within healthcare in major areas such as the evaluation of treatment effectiveness, management of healthcare, customer relationship management, and the detection of fraud and abuse. It also gives an illustrative example of a healthcare data mining application involving the identification of risk factors associated with the onset of diabetes [6].

**4.7 P. Samuel Kirubakaran, S.Uma, Donbyntalang Dewkhaid, Selva Kumar (2015)**, They presents the m-Healthcare system which is envisioned as an important application of pervasive computing to improve health care quality and to save lives. Implantable body sensor nodes and smart phones are utilized to provide remote healthcare monitoring to people who have chronic medical conditions such as diabetes and heart disease. Medical users receive the high-quality healthcare monitoring from medical professionals anytime and anywhere [7].

**4.8 Dr. R. Hariharan (2016)**, analyses the health status of the rural women in India. The review analysed the socio-cultural constraints of women, health needs, widespread women's ignorance, women's rights, women's access to available resources, nutritional stress, fertility and reproductive health, gender discrimination and awareness of reproductive health infections [8].

**4.9 R. Naveen Kumar, M. Anand Kumar (2016)**, their work presents a brief preamble of the techniques that are currently used in medical field followed by virtues and demerit of the existing techniques. Their study

also highlights applications, challenges and future issues of Data Mining in healthcare and provides the recommendations for appropriate selection of available Data Mining techniques [9].

**4.10 Minerva Panda, Syed Mohd Ali, Sanjog Kumar Panda (2017)**, they focus on the development of a mobile or web application, through which patients send their symptomatic query to the doctors through a server. The mobile application will be equipped with first aid instructions, according to the nature and severity of the symptoms, either the patients are directed to respective departments or given emergency help for further treatment. [10]

**Table-1.0: Comparative Table of different Data Mining Approaches/Techniques/Tools used in Healthcare Domain**

S. No.	Research Title	Authors / Year	Data Mining Approaches/ Techniques/ Tools/ Algo discussed in the paper	Key-points	Results	Limitations/ Challenges
1.	Data Mining Applications In Healthcare Sector: A Study	M. Durairaj, V. Ranjani (2013)	Rough Set Theory & ANN	compare the different data mining application in the healthcare sector for extracting useful information	The comparison study shows that data mining techniques in all the health care applications give a more encouraging level of accuracy like 97.77% for cancer prediction and around 70 % for estimating the success rate of IVF treatment.	Exploring knowledge from the medical data is such a risk task as the data found are noisy, irrelevant and massive too.
2.	Prediction and Decision Making in Health Care using Data Mining	Boris Milovic, Milan Milovic (2012)	Genetic Algorithm, ANN, Decision Tree	Main obstacle is that almost all diagnoses and treatments in medicine are imprecise and subjected to error rates.	To predict future requests, needs, desires, and conditions of the patients and to make adequate and optimal decisions about their treatments	available raw medical data are widely distributed, different and voluminous by nature
3.	A survey on Data Mining approaches for Healthcare	Divya Tomar and Sonali Agarwal (2013)	classification, clustering, association, regression	In MDS there are 300 questions which are answered by the patients at check-in time	They conclude that there is no single DM techniques which give consistent results for all types of healthcare data. The performance of DM techniques depends on the type of dataset that we have taken for doing experiment	difficult to acquire the precise and complete healthcare data. It is complex and heterogeneous in nature
4.	A Text Mining Approach to Automated Healthcare for the Masses	Vishnu S. Pendyala, Yi Fang, JoAnne Holliday and Ali	Labeling, QA Interface, KNN, NLP	work on automating medical diagnosis using text mining techniques and include	presented our vision of how machines can take-on a very essential function of life support and proposed an automated health monitoring, diagnosis & regulation	the key challenges that still need to be overcome privacy & security, cost, data set and

		Zalzala (2014)		some initial results		human expertise
5.	Applying Data Mining Techniques to Standardized Electronic Health Records for Decision Support	Shivani Batra, Shelly Sachdeva, Hem Jyotsana Parashar and Pulkit Mehndiratta (2013)	Classification, Clustering, Association Rule	to identify the need of applying data mining techniques to standardized electronic healthcare records	identifies and extracting useful information from standardized health care database. Mined knowledge may be used to provide better medication and early detection of disease & reducing the cost of treatment.	Challenges in data mining to make standardized database to store incorrect data, kinds of user, granularity of data and volatility.
6.	Data Mining Applications in Healthcare	Hian Chye Koh and Gerald Tan (2015)	Classification table, Decision Table, chi square	Discuss about data mining and its applications within healthcare such as the evaluation of treatment effectiveness, mgt of healthcare, customer relationship mgt, and detection of fraud and risk.	developed to better identify and track chronic disease states and high-risk patients, design appropriate interventions, and reduce the number of hospital admissions and claims	accessibility of data, raw inputs and other data problems may arise, these include missing, corrupted, inconsistent, or non-standardized data, such as pieces of information recorded in different formats in different data sources.
7.	An Automated Health Care Computing Model for Continuous Monitoring of Patients for Immediate Medical Care during Emergency	P. Samuel Kirubakaran, S.Uma, Donbyntalang Dewkhaid, Selva Kumar (2015),	Pervasive computing, Implantable body sensor, Ontology learning, Software Engineering	To deliver health care services to meet the health needs of target populations. To provide increasing access and decreasing unnecessary variations in proper care.	make m-Healthcare system which is envisioned as an important application of pervasive computing to improve health care quality and to save lives	m-Healthcare still faces many challenges including information security and privacy preservation.
8.	Health Status of Rural Women in India: An Overview of Literatures	Dr. R. Hariharan (2016),	Decision Making, Learning outcomes	women faces various unique health issues as compared to male, there is a need for more specific and combined	analysed the socio-cultural constraints of women, health needs, widespread women's ignorance, women's rights, women's access to available resources, nutritional stress, fertility and reproductive health, gender discrimination	suggests the researchers in the field of women health to bring various research for safeguarding the women health status as whole

				research on women health status	and awareness of reproductive health infections	
9.	Medical Data Mining Techniques for Health Care Systems	R. Naveen Kumar, M. Anand Kumar (2016),	classification, clustering, association, regression	presents a brief preamble of the techniques that are currently used in medical field followed by virtues and demerit of the existing techniques	detection of disease diagnosis by performing classification or clustering techniques. But still there is a low concentration on the quality of the dataset, when it is incomplete.	The preprocessing of dataset is less concentrated in the existing work, which leads to major setback in the overall process.
10.	Big Data in Health Care: A Mobile Based Solution	Minerva Panda, Syed Mohd Ali, Sanjog Kumar Panda (2017)	Big Data, Health Analytics	unified model can serve as a data collection, delivery as well as an analytic tool in the healthcare domain.	developed a mobile/web application, through which patients sends their symptomatic query to the doctors through a server and patient can take help during emergency	key challenges posed to in healthcare system that are shortage of human resources, accessibility of healthcare infrastructure, affordability of healthcare services especially for the rural populations

*Sources: Authors Compilation*

**Table-2.0: Comparative Table on different Data Mining Algorithms & their Learning Types used in Healthcare Domain**

S. No.	Algorithm/ Tool	Learning Type	Behavior	Task Performed	Evaluation Criteria	Performance depends on
1.	Classification	Supervised	Predictive	Identifying Class	Accuracy	Correctness of Training Data
2.	Clustering	Unsupervised	Descriptive	Grouping Data	Accuracy	Efficiency
3.	Regression	Supervised	Predictive	Map a data item	Evaluation	Type of function (linear, logistic etc)
4.	Association Rule	Unsupervised	Descriptive	Identifying correlated attributes	Efficiency	Support and confidence value
5.	Prediction	Supervised	Predictive	Predicting future dataset	Accuracy	Past and current data
6.	Summarization	Unsupervised	Descriptive	Maps data into subset	Accuracy	Derived data from data
7.	Time Series Analysis	Supervised	Predictive	visualize time series	Forecasting	Value of an attribute
8.	Sequence Discovery	Unsupervised	Descriptive	Sequential pattern in data	Efficiency	Based on time

*Sources: Authors Compilation*

## 5. Conclusion:

After reviewing the different studies, the researcher found that there are many constraint to explore the knowledge and take effective decision from medical databases because the available raw medical data are widely distributed, non-standardized, noise, corrupted, inconsistent, missing, voluminous by nature and recorded in different file formats. This review puts focus on challenges to using data mining on healthcare domain, it shows the important concern about information or data security & privacy preservation, cost, data set, reduce error and human expertise. Our critical study indicates that there are scopes to overcome these constraints and limitations. The limitations of the study clearly mentioned in the above comparative table 1.0 and 2.0, which shows the scope for further research.

## References:

- [1] M. Durairaj, V. Ranjani (2013), Data Mining Applications in Healthcare: A Study, International Journal of Scientific & Technology Research, volume 2, issue 10, October 2013, ISSN 2277-8616
- [2] Boris Milovic, Milan Milovic (2012), Prediction and Decision Making in Health Care using Data Mining, International Journal of Public Health Science (IJPHS), Vol. 1, No. 2, December 2012, pp. 69~78, ISSN: 2252-8806
- [3] Divya Tomar and Sonali Agarwal (2013), A survey on Data Mining approaches for Healthcare, International Journal of Bio-Science and Bio-Technology Vol.5, No.5, pp. 241-266 <http://dx.doi.org/10.14257/ijbsbt.2013.5.5.25>
- [4] Vishnu S. Pendyala, Yi Fang, JoAnne Holliday and Ali Zalzal (2014), A Text Mining Approach to Automated Healthcare for the Masses, IEEE, Global Humanitarian Technology Conference, ISBN 978-1-4799-7193-0/14/
- [5] Shivani Batra, Shelly Sachdeva, Hem Jyotsana Parashar and Pulkit Mehndiratta (2013), Applying Data Mining Techniques to Standardized Electronic Health Records for Decision Support, IEEE, Sixth International Conference on Contemporary Computing (IC3), ISBN 978-1-4799-0192-0
- [6] Hian Chye Koh and Gerald Tan (2015), Data Mining Applications in Healthcare, Journal of Healthcare Information Management — Vol. 19, No. 2
- [7] P. Samuel Kirubakaran, S.Uma, Donbyntalang Dewkhaid, Selva Kumar (2015). An Automated Health Care Computing Model for Continuous Monitoring of Patients for Immediate Medical Care during Emergency. International Journal of Computer Science and Information Technologies (IJCSIT), Vol. 6 (2) , 2015, 1307-1311. ISSN 0975-9646
- [8] Dr. R. Hariharan (2016), Health Status of Rural Women in India: An Overview of Literatures, International Journal of Research in Economics and Social Sciences (IJRESS), Vol. 6, Issue 8, August , 2016, pp. 109~119, ISSN(o): 2249-7382 | Impact Factor: 6.225, | Thomson Reuters ID: L-5236-2015
- [9] R. Naveen Kumar, M. Anand Kumar (2016), Medical Data Mining Techniques for Health Care Systems. International Journal of Engineering Science and Computing (IJESC), ISSN 2321 3361, DOI 10.4010/2016.811.
- [10] Minerva Panda, Syed Mohd Ali, Sanjog Kumar Panda (2017), Big Data in Health Care: A Mobile Based Solution, IEEE International Conference On Big Data Analytics and computational Intelligence (ICBDACI), 978-1-5090-6399-4/17/
- [11] Sanjay Jaiswal (2015), Rural Healthcare: Towards a Healthy Rural India, Gram Vani: The Voice of Village India, <http://www.gramvaani.org/?p=1629>
- [12] Margaret H. Dunham (2009), Data Mining: Introductory and Advanced Topics, Pearson Education, ISBN 978-81-7758-785-2