Data Mining in Health Care: Application Perspective

A. S. Shanthakumari and R. Jayakarthik

Abstract In order to attract the attention of doctors, nurse practitioners, clinical pharmacist, physician assistants and scientists alike, the rising health sector is producing a massive quantity of patient personal details and imbursements. The aims of this paper are to compare the different techniques, approaches and tools and also to measure their effectiveness in the healthcare sector. The main objective of the data mining application is to convert data into facts, text or number of applications that have been refined into knowledge by a computer. The purpose of applying data mining is to devise a programmed tool to identify and inaugurate relevant healthcare information in the healthcare industry. The researcher aims to study different types and challenges of data mining applications in the healthcare industries. Lastly, it also shows the past data mining techniques and its implementation methods.

Keywords Data mining · Application · Health care · Algorithms · Diseases

1 Introduction

Some of the papers have recently discussed various aspects of data mining's use in health care. In this paper the researcher used data mining algorithms to present a summary on healthcare analytics. In the current scenario of application of data mining in health care, create and gather high volumes of information that assist to give some interesting patterns from electronic systems that will protect medical records and enable quick detection of infectious diseases. Data mining facilitates healthcare

A. S. Shanthakumari (✉) · R. Jayakarthik
Department of Computer Science, VISTAS, Chennai, Tamil Nadu, India

A. Ramu et al. (eds.), *International Conference on Computing, Communication, Electrical and Biomedical Systems*, EAI/Springer Innovations in Communication and Computing, https://doi.org/10.1007/978-3-030-86165-0_13

sectors to forecast trends in the patient's health state by building links between apparently disparate information. The raw data from healthcare sectors needs to be stored, and their combination allows the formation of a connected medical information system [1].

2 Literature Survey

Durairaj and Ranjani [2] in his article found that the raw data from healthcare sectors are huge and assorted. The data needs to be collected, and the extracted patterns can be utilized by healthcare practitioners to construct forecasts and locate treatments for patients in healthcare sectors.

Ţăranu [3] suggested an innovative perspective on decision-making strategies for the health data to give examples in the healthcare industry through data mining techniques. The theoretical framework of data discovery in databases concluded that the problems in the healthcare sector are identified by data mining applications.

Karegar et al. [4] constructed an automated tool for relevant healthcare information of various types of data mining applications and to minimise the complexity of the healthcare data transactions. The past data mining techniques and its function tools for healthcare organisations are also discussed.

Islam et al. [5] reviewed the numerous papers in this area concerned in terms of algorithms, methods and performance. Their analysis paper summarised the papers that the disciplines, pattern, errands and methods attempted inline.

Cifci and Hussain [6] were motivated by a lack of a coherent and structured narrative to frame a summary of the literature on this specific issue. From 2005 to 2016, they performed a sample search. They examined that the current literature explores analytics in decision-making in clinical and directorial. They were extracted through sub-areas of health care, DM techniques and forms of analytics to provide potential future directions.

Mercy Beulah et al. [7] revealed in their paper that data mining is one of the important motivating spaces for investigation that is mounting gradually standard in the healthcare sector. DM plays a vital role in revealing the innovative emerging trends related to these circumstances. Their survey highlighted some of the applications and the opportunity issues of DM in medical sectors. They also provided a picture of a database of the present situation held in the healthcare sector.

3 Data Mining

Data mining is a method by which patterns are extracted from data. With prevalent usage and unpredictable growth in database sizes, the sectors are confronted with the question of using such huge quantities of data in companies. Knowledge

discovery in databases is an important extraction of embedded data that gives potentially useful information from the data [8, 9].

4 Methodology

The degree of precision is claimed by the applications. Sterility requires complex methodologies to conclude the outcomes at the end of IVF (in vitro fertilization) treatment specific procedures. In this paper the researcher compares the different data mining techniques to the same standard to estimate IVF treatment success rate. This comparative study may be useful in which data mining method offers a reliable level of knowledge from data from health care.

5 IVF Success Rate Prediction Comparison

The data mining derives the conclusions from the data and presumes the level of precision in different healthcare issues. Figure 1 shows data mining applications in health care. Table 1 shows data mining applications in health care. The list for assessment of medical conditions is:

- Heart disease.
- Cancer.
- Tuberculosis.
- Diabetes.
- Kidney dialysis.
- Dengue.

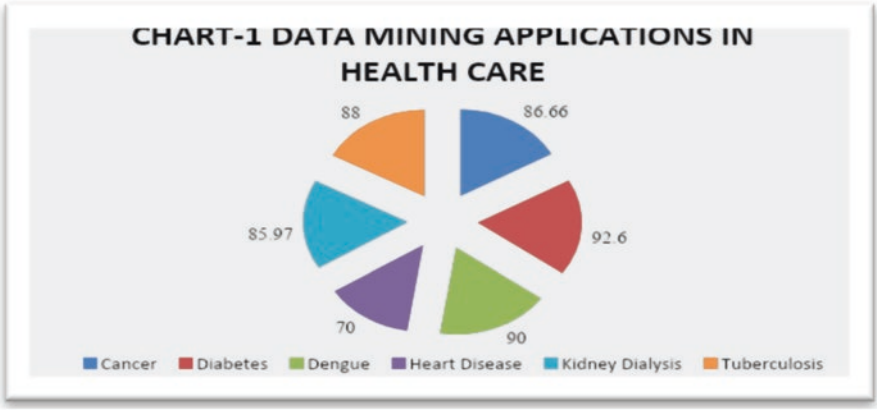


Fig. 1 Data mining applications in health care

Table 1 Data mining applications in health care

S. no.	Diseases	Tool of data mining	Technique	Algorithm	Traditional method	Accuracy level (%) from DM application
1.	Cancer	WEKA	Classification	Decision table	–	86.66
2.	Diabetes	ANN	Classification	C4.5	Neural network	92.6
3.	Dengue	SPSS Modeller	–	C5.0	Statistics	90
4.	Heart disease	NCC2	Classification	Naive	Probability	70
5.	Kidney dialysis	RST	Classification	Decision-making	Statistics	85.97
6.	Tuberculosis	WEKA	Naive	KNN	Probability, statistics	88

Table 2 Estimates of the rate of success of the operation and presents the outcome

	Predicted			
Actual		**Success**	**Unsuccess**	
	Success	17	4	0.80
	Unsuccess	26	10	0.277
		0.3953	0.7142	0.4736

6 Application of Rough Set Theory for Medical Informatics Data Analysis

Development of the rough set theory is used in medicinal data evaluations. Raw collection is used to lower the attributes, without losing the relationship of the initial. The fertilization data to construct the optimal editing is set without having any effect on the original relationship. Table 2 shows the estimates of the rate of success of the operation and presents the outcome.

Comparison is made between expected and real outputs. It also indicates that the performance rate after attribute reduction is 47%.

7 Artificial Neural Network in Classification and Prediction

The network being modelled is evaluated, trained and confirmed using patient sample IVF data. Finally, it measures the rate of performance between expected outcome and actual production. Table 3 accounts for a success rate foreseen by ANN.

Table 3 IVF success rate predicted by ANN

Performance	Desired output	Actual network output
MSE	0.209	0.2128
NMSE	1.164	1.1830
MAE	0.231	0.2578
Min. Abs error	9.908	6.6604
Max. Abs error	1.015	0.9988
R	0.0498	0.4980
Percent correct	73.07	75

Table 4 Performance of IVF success rate prediction using hybrid technique

Performance	Unsuccess of treatment	Success of treatment
MSE	0.093	0.110
NMSE	0.379	0.451
MAE	0.143	0.192
Min. Abs error	0.003	0.006
Max. Abs error	1.056	1.056
R	0.789	0.789
Percent correct	89.230	91.837

8 Modelling Neural Networks and Rough Sets for Analysing Medical Data

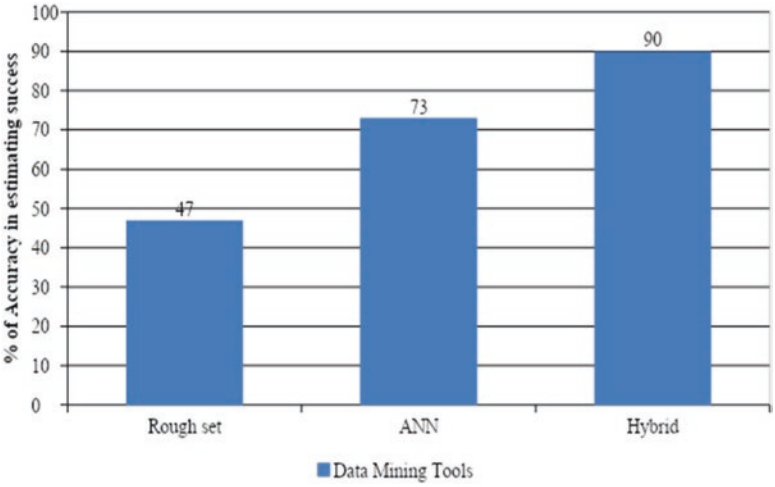
In the application of rough set method, the two forms of rules destinism and non-destinism are affected. The presentation of artificial neural network's collective technique and rough set theory is explained in Table 4.

In Table 4, this hybrid method predicts the accuracy of ANN and RST collectively. Collective rough collection and artificial neural network implementation yields better outcomes when opposed to other techniques. Table 5 shows comparison between two different data mining applications.

9 Conclusion

It is a difficult task to hypothesize disorders using data mining tools, but it significantly minimizes manual workers and maximizes prediction value. Improving efficient data mining software minimizes the expense of HR and experience and time constraints. Exploring medicinal data knowledge is such a dangerous activity as the collected data is always chaotic, immaterial and enormous, such as 86.66% for cancer prediction and about 70% for estimating the rate of success of IVF treatment.

Table 5 Comparison between two different data mining applications



References

1. Canlas, R.D.: Data Mining in Healthcare: Current Applications and Issues. School of Information Systems & Management, Carnegie Mellon University, Australia (2009)
2. Durairaj, M., Ranjani, V.: Data mining applications in healthcare sector: a study. Int. J. Sci. Technol. Res. **2**(10), 29–35 (2013)
3. Ţăranu, I.: Data mining in healthcare: decision making and precision. Database Syst. J. **6**(4), 33–40 (2016)
4. Karegar, M., Isazadeh, A., Fartash, F., Saderi, T., Navin, A.H.: Data-mining by probability-based patterns. In: ITI 2008-30th International Conference on Information Technology Interfaces, pp. 353–360 (2008)
5. Islam, M.S., Hasan, M.M., Wang, X., Germack, H.D.: A systematic review on healthcare analytics: application and theoretical perspective of data mining. Healthcare. **6**(2), 1–43 (2018)
6. Cifci, M.A., Hussain, S.: Data mining usage and applications in health services. Int. J. Inform. Visualiz. **2**(4), 225–231 (2018)
7. Mercy Beulah, E., Nirmala Sugirtha Rajini, S., Rajkumar, N.: Application of data mining in healthcare: a survey. Asian J. Microbiol. Biotechnol. Environ. Sci., **18**(4), 999–1001 (2016)
8. Datta, D., Mishra, S., Rajest, S.S.: Quantification of tolerance limits of engineering system using uncertainty modeling for sustainable energy. Int. J. Intell. Netw. **1**, 1–8 (2020)
9. Roski, J., Bo-Linn, G.W., Andrews, T.A.: Creating value in health care through big data: opportunities and policy implications. Health Aff. **33**(7), 1115–1122 (2014)