INTELLIGENT MEDICAL DATA ANALYTICS: A CRUCIAL CHALLENGES & REVIEWS

Article *in* The Journal of critical illness · July 2020

DOI: 10.31838/jcr.07.06.09

1 author:

Kalpana Chelladurai
Vels University
4 PUBLICATIONS **0** CITATIONS

SEE PROFILE

Journal of Critical Reviews

ISSN- 2394-5125 Vol 7, Issue 6, 2020

Review Article

INTELLIGENT MEDICAL DATA ANALYTICS: A CRUCIAL CHALLENGES & REVIEWS

C. Kalpana[1], Dr.B. Booba[2]

[1]Research Scholar, Dept. of CSE VISTAS, Vel's University, Chennai. rkalpz@gmail.com
[2]Professor, Dept. of CSE VISTAS, Vel's University, Chennai. boobarajashekar@gmail.com

Received: 09.02.2020 *Revised: 11.03.2020* *Accepted: 13.04.2020*

Abstract

In the present world, the healthcare industries demand and costs may be increasing due the high population. It is a huge challenge for healthcare industries to analyse and classify disease. The available doctors cannot cover the all patients. The healthcare or medical informations are investigating and information management became a huge challenge because the complexity of problems. Data mining and data warehousing for Biological or medical problems is one of the most challenging issues in the real world. Human medical information are the mainly difficult of all biological information to classify to analyze the problem. In current senario, a number of peer-reviewed research and reviw articles have deal with different aspects of healtcare data analytics applications. However, the shortfall of extensive and methodical clarification motivated us to build a extensive literature review and chalanges on this theme. In this paper, a systematic review of the comprehensive literature on healthcare intelligent data analytics.

Keywords: Medical Data Analysis, Data Development, Intelligent Decision-making.

DOI: http://dx.doi.org/10.31838/jcr.07.06.09

INTRODUCTION

In the recent era of intelligent medical data analysis, data development, variety of data formats and minimum time for intelligent decision-making are the important challenges for healthcare industries are facing. Information is a major important asset of the data mining and data warehousing organization. Noramlly the organizations are store the all information like aquisation of data from patients, disease identification, treatments and its transaction related data for its future use in analysis purpose. The digital transformation provided comparatively inexpensive information storage devices helped to store all data in the form of huge data sorage systems. Internet is an ideal option for lot of opportunities to transact and deliver the information to the various organization to do business and analysis. Data mining and wareousing is a procedure that uses a diversity of data analysis methodologies to determine various patterns and variety of relationships in information that may be employed to construct a valid predictions. The datamining process is to mine the useful data through the available dataset. Based on the various steps and various clusters to find the useful data for valuable decision making [1-3].

Normally the Data mining find out the databases to search hidden patterns and evaluavate the information for better decision support through intelligent algorithm. In the medical investigational approach, the Swarm intelligence and data analytics methods are implemnted to extract resonable information from huge collection of patients information for better decision making process.

The Classification alogrithms, clustering methods, prediction methodologies and sequence detection types are the various classification methos to solve the data mining problems. Statistical analysis, machine learning and pattern recognition algorithms, decision trees, neural network, rule induction, fuzzy sets, bayesian networks, rough sets algorithms and genetic algorithms are the various techniques used in data mining process. Figure 1shows the Medical data analytics framework. Figure 2 shows the Classification of Healthcare/ Medical Data Analytics and its applications.

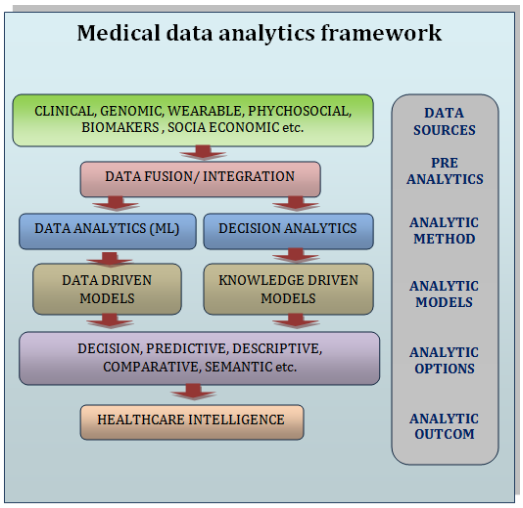

Fig. 1: Medical Data Analytics Framework

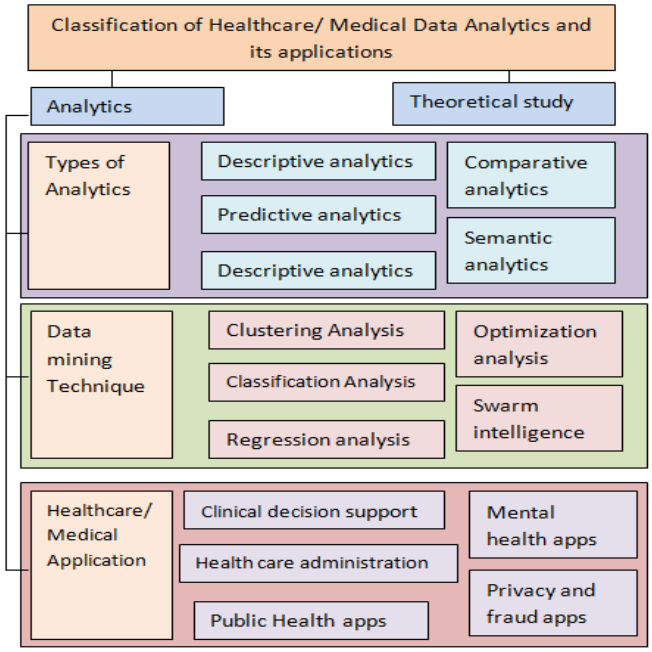

Fig. 2: Classification of Healthcare/ Medical Data Analytics and its Applications

Challenges in Healthcare/Medical System

There is a numberous challenges that makes it too difficult for efficient implemenation of health care system to its fullest extent. Clinical or medical data such as disease information, patient history, prliminary vital signs, health progress notes, health monitoring imformation and the diagnostic tests results of are stored in the database. Clinical information is accessed, monitored and maintained by the doctors, nurses, and other clinical staff. Based on the information, the doctors track the patient care and to communicate treatment methods and plans to the clinical team members.

In this view, the good healthcare management for increasing the efficiency to the society. Information is crucial for collecting and transacting to all management levels of the health services. It is crucial challenges for patient/client management system, for health unit management, as well as for health system planning and management [3-7].

- The new challenging issues of high-dimensional data are a very complicated problem because the system representation includes many attributes and its related values connected with sparsity and noise problems.
- The feature extraction methods are used to reduce the features in to lower dimensional feature reduced space. This feature extraction technique each and every dimension of features is compatible to the grouping of original features. The feature extraction techniques classified in to two major types such as linear or non-linear.
- The feature selection technique is the important activity of finding out a group of significant feature attributes in the medical system. The feature selection solving problems includes optimization methods, wrapper techniques, hybrid algorithms, filtering and statistical methods.
- The Health care and Medical information go through frequently from abnormality and complex problem.

- Missing medical information is to be big threat by either unproductive information compilation or lack of observation from the hospital to the document. In data acquisition issues, most of the patients are unable to check with proper manner. In observation with document problem, a patient health information is comes under negative results.
- Health information suffers from a noisy problem; due to coding inaccuracies, inconsistent naming, conventional techniques, etc. Several investigators tend to implement Meta heuristic/machine learning algorithms to find more robust depiction and to solve the noisy problem.
- The effectiveness of the automated diagnosis system, intelligent diagnosis system and clinical decision system demonstration is very less.
- The medical information system is partial because of too many missing attributes of the machine learning system. The medical information is a problem must be investigated by optimized machine learning algorithms.
- The healthcare information is normally very noisy because of few factors such as coding inaccuracies or error in coding, input error etc.
- Patient education and medical education with past histories of patient reports integration is a major problem in the database access.

LITERATURE REVIEW

In order to investigate the management information system and knowledge about implementing mehodoloies are at minimum levels of hierarchy in the healthcare system. The table 1 focus on the usage and review of Health Information system.

Author	Methods	Remarks
Chistianini and Shawe Taylor [8]	Support Vector Machine	Devoping hyper plane by Support Vector Machine (SVM) is to classify the data points of the medical system.
Sambasivan et al [9]	Decision Support System	Extracting and evaluating the past patient details and information are stored in the database. Then predictive methods implemented through forecasting, reducing time and costs.
Aditya B. Patelet al [10]	Map Reduce programming	All reports and the experimental work stored in the database. Then findout the optimal solutions using Hadoop cluster.
Mukherjee, A. et al [11]	Map Reduce model	Map-Reduce Frame work is implemented in Big data analytics to classify and analysis of large amount of data with hidden patterns.
Khan et al., [12]	decision tree	The decision tree is used for predicting, analyzing and classifying the breast cancer patient in the large database system.
E.Avci et al., [13]	genetic support vector machines	It is proposed that the heart valve disease intelligent system using genetic support vector machines (GVSM) for better investigational purpose.
Garlasu D et al [14]	Hadoop technology	The Hadoop technology is implemneted in the process & storage purpose.
Liu et al., [15]	BBN	The health data is clearly analysed and making a better decision support system using BBN.
Er et al., [16]	Artificial Neural Network	It ia proposed to build a system using Artificial Neural Network (ANN) for efficieant analyzing and comparison of chest diseases.
Jena et al., [17]	Linear Discriminate Analysis	Linear Discriminate Analysis(LDA) method is used for early warning system in chronic disease.
Moon et al., [18]	decision tree algorithm	consumption of alcohol and psychological distress problems are analysed by using decision tree algorithm .
Shouman et al., [19]	K-NN classifier	heart disease analysis and diagnosis of heart disease system by using K-NN classifier .
Abdi et al.,[20]	PSO based SVM	PSO with SVM model is implemented to investigate the erythemato -squamous diseases
Zuoa et al.,[21]	adaptive Fuzzy K-NN	Adaptive Fuzzy K-NN methodologies is used to identify the Parkinson disease.
Chen et al.,[22]	hierarchical K-means	It is proposed that hierarchical K-means technique is used for better analyzing large micro-array data.
Chipman et al., [23]	hybrid hierarchical clustering approach	hybrid hierarchical clustering method is used toinvetigate the microarray data.
R. Agarwal et al [24]	Apriori Algorithm	The Apriori algorithm and association rules (sets of transactions) implemented in medical database.
Patil et al., [25]	association rules	type-2 diabetes patient classification is implemented using association rules.
Ying et al., [26]	fuzzy recognition-prime decision	The relationship between drugs is analysed with proposed data mining association approach and fuzzy recognition-prime decision (RPD) model.
Kai et al., [27]	clinical decision support system	clinical decision support system for healthcare workers to identify the noncommunicable diseases.
Cervantes et al., [28]	binary method	The simplest is the binary method is used to classify the patient data with significant attributes.
Tanet al., [29]	data clustering analysis	The data clustering analysis technique is implemented in medical data into variois clusters.
Murphy, [30]	Similarity measure	clustering analysis and machine learning is used to classify the grouping similar objects.
Shi, [31]	brain storm	Health care analysis and classification is done by using the the brain storm

	optimization	optimization algorithm and clustering analysis.
Nazmul Siddique1, Hojjat Adeli, [32]	nature-inspired searching techniques	The nature-inspired searching techniques and aSIalgorithm are used for the processing and analysing data.
Ficici, [33]	coevolution	As a general Swarm intelligence principleis used to find the fitness of a solution with in the computational resources along with time and/or space.
Kennedy and Eberhart, [34]	PSO	particle swarm optimization algorithm is used to solve the continuous optimization problems
Dorigo and Stutzle, [35]	ACO	The Ant Colony Optimization algorithm is used find out the short route of the signal transmission in the network.
Martens et al., [36]	data mining techniques	swarm individuals is move through a solution search space and find the optimum solution for the data mining task.
Abrahamet al., [37]	clustering with differential evolution	data mining concepts is implemnted to find out the single objective for document clustering
Ahmadi et al., [38]	particle swarm	multiobjective problems soloved by using particle swarm
Slaney and Casey,[39]	locality sensitive hashing algorithm	high-dimensional large database problem is solved by locality sensitive hashing algorithm and to identify the similar entries.
Lu et al., 2011 [40]	Particle swarm optimizer	PSO is used to clustering high-dimensional data for the different problem in the datamining task.

CONCLUSION

The healthcare data analytics area is a new and latest topic. The large volume of medical sector information is also incresing day by day. Investigating the high-dimensional health information and the multiple aim of the healtcare information are also vital role in solving big data problems. In this paper, various research articles are reviewed and the motivation of using variety of data mining and optimization algorithm in healthcare and related area. It is also presented the various healthcare and medical data with Artificial intelligence are analysed and surveyed with major disease.

REFERENCES

1. PC Roy, SR Abidi, SSR Abidi, "Possibilistic activity recognition with uncertain observations to support medication adherence in anassisted ambient living setting", Knowl Based System. Vol.133, pp. 156-173, 2017.
2. SE Gabriel, SL Normand, "Getting the methods right—the founda-tion of patient-centered outcomes research", The New England Journal of Medicine, Vol. 367(9), pp.787-790, 2012.
3. F Jiang, Y Jiang, H Zhi, "Artificial intelligence in healthcare:past, present and future", Stroke and Vascular Neurology journal. Vol. 2(4), pp. 230-243, 2017.
4. TB Murdoch, AS Detsky, "The inevitable application of big data to healthcare", JAMA., Vol. 309(13) pp. 1351-1352, 2013.
5. Z Obermeyer, EJ Emanuel. Predicting the future—big data, machine learning, and clinical medicine, The New England Journal of Medicine, 2016; 375(13):1216-1219.
6. H Yang, OE Kundakcioglu, J Li, T Wu, "Healthcare intelligence-turning data into knowledge", IEEE Intelligent Systems, Vol. 29(3) pp.54-68, 2014.
7. N Kureshi, SSR Abidi, CA, "Blouin predictive model for perso-nalized therapeutic interventions in non-small cell lung cancer", IEEE Journal of Biomedical and Health Informatics, Vol. 20(1), pp.424-431, 2016.
8. N Chistianini and J Shawe-Taylor, "An Introduction to Support Vector Machines, and other kernel-based learning methods", Cambridge University Press, 2000.
9. Sambasivan, Murali; Pouyan Esmaeilzadeh, Naresh Kumar; Hossein Nezakati, "Intention to adopt clinical decision support systems in a developing country: effect of Physician's perceived professional autonomy, involvement and belief: a cross-sectional study". BMC Medical Informatics and Decision Making, Vol.12, pp.142–150, 2012.
10. B Aditya. Patel, Manashvi Birla, Ushma Nair, "Addressing Big Data Problem Using Hadoop and Map Reduce" in IEEE International Conference on Data Mining and Network Technologies , 2012.
11. A Mukherjee, J Datta, R Jorapur, R Singhvi, S Haloi, W Akram, "Shared disk big data analytics with Apache Hadoop". IEEE 19th international conference In High Performance computing (HiPC), Washington, DC, USA ,pp. 1–6, (2012).
12. MU Khan, JP Choi, H Shin, M Kim, "Predicting Breast Cancer Survivability Using Fuzzy Decision Trees for Personalized Healthcare", 30th Annual International IEEE EMBS Conference Vancouver, British Columbia, Canada, 2008.
13. Avci, "A new intelligent diagnosis system for the heart valve diseases by using genetic-SVM classifier", Expert Systems with Applications, Elsevier, vol. 36, pp. 10618-10626, 2009.
14. D Garlasu, Virginia Sandulescu, Ionela Halcu, Giorgian Neculoiu, Oana Grigoriu, Mariana Marinescu, Viorel Marinescu, "A big data implementation based on grid computing", 11th IEEE international conference on RoEduNet, 2013.
15. KFR Liu, C F Lu, "BBN-Based Decision Support for Health Risk Analysis", Fifth International Joint Conference on INC, IMS and IDC, 2009.
16. O Er, N Yumusakc, F Temurtas, "Chest diseases diagnosis using artificial neural networks", Expert Systems with Appls, vol. 37, pp. 7648-7655, 2010.
17. H Jena, C C Wang, B C Jiangc, Y H Chub and M S Chen, "Application of classification techniques on development an early-warning systemfor chronic illnesses", Expert Systems with Applications, vol. 39, , pp. 8852-8858, 2012.
18. SS Moon, SY Kang, W Jitpitaklert, SB Kim, "Decision tree models for characterizing smoking patterns of older adults", Expert Systems with Applications, Elsevier, vol. 39, pp. 445-451, 2012.
19. M Shouman, T Turner and R Stocker, "Applying K-Nearest Neighbour in Diagnosing Heart Disease Patients", International Conference on Knowledge Discovery (ICKD-2012), 2012.
20. M J Abdi, D Giveki, "Automatic detection of erythemato-squamous diseases using PSO–SVM based on association rules", Engineering Applications of Artificial Intelligence, vol. 26, pp. 603-608, 2013.
21. WL Zuoa, ZY Wanga, T Liua, HL Chenc, "Effective detection of Parkinson's disease using an adaptive fuzzy k-nearest neighbor approach", Biomedical Signal Processing and Control, Elsevier, Vol. 8(4), pp. 364-373, 2013.
22. TS Chen, TH Tsai, YT Chen, CC Lin, RC Chen, SY Li, H Y Chen, "A Combined K-Means and Hierarchical Clustering Method for improving the Clustering Efficiency of Microarray", Proceedings of 2005 International Symposium on Intelligent Signal Processing and Communication Systems, 2005.

23. Chipman, R Tibshirani, "Hybrid hierarchical clustering with applications to microarray data", Biostatistics, vol. 7, no. 2, pp. 286-301, 2006.
24. R Agrawal, T Imielinski, A Swami, "Mining association rules between sets of items in large databases", Proceedings of the ACM SIGMOD International Conference on the Management of Data. ACM, Washington DC, pp. 207–216, 1993.
25. M Patil, R C Joshi, D Toshniwal,"Association rule for classification of type -2 diabetic patients", Second International Conference on Machine Learning and Computing, 2010.
26. J Yanqing, H Ying, J Tran, P Dews, A Mansour R Michael Massanari, "Mining Infrequent Causal Associations in Electronic Health Databases", 11th IEEE International Conference on Data Mining Workshops, 2011.
27. Eiko Kai1, Andrew Rebeiro-Hargrave, Sozo Inoue, Yasunobu Nohara, Rafiqul Islam Maruf, Naoki Nakashima, Ashir Ahmed, "Enpowering the healthcare worker using the Portable Health Clinic", IEEE Transactions, DOI 10.1109/AINA.2014.108, 2014.
28. A Cervantes, I M Galván, P Isasi, "AMPSO: a new particle swarm method fornearest neighborhood classification", IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, Vol. 39 (5), pp. 1082-1091, 2009.
29. PN Tan, M Steinbach, V Kumar, Introduction to Data Mining, Addison Wesley, Boston, MA, 2005.
30. KP Murphy, Machine Learning: A Probabilistic Perspective, Adaptive Computation and Machine Learning Series, The MIT Press, Cambridge, MA, 2012.
31. Y Shi, "Brain storm optimization algorithm", Advances in Swarm Intelligence, Vol. 6728, lecture notes in computer science, Springer, Berlin and Heidelberg, pp. 303-309, 2011.
32. Nazmul Siddique1, Hojjat Adeli, "Nature Inspired Computing: An Overview and Some FutureDirections", Cognitive Computation, Vol 7, pp.706–714, 2015.
33. S G Ficici, "Monotonic solution concepts in coevolution", Genetic and EvolutionaryComputation Conference (GECCO 2005), June, pp. 499-506, 2005.
34. J Kennedy, R Eberhart, "Particle swarm optimization", Proceedings of IEEE International Conference on Neural Networks (ICNN), pp. 1942-1948, 1995.
35. M Dorigo, T Stu¨tzle, The ant colony optimization metaheuristic: Algorithms, applications andadvances. Handbook of Metaheuristics, vol. 57 of International Series in Operations Research & Management Science (pp. 251–285). Norwell, MA, Kluwer Academic Publishers
36. D Martens, B Baesens, and T Fawcett, ,"Editorial survey: swarm intelligence for data mining", Machine Learning, Vol. 82 No. 1, pp. 1-42, 2011.
37. A Abraham, S Das, and A Konar,),"Document clustering using differential evolution",Proceedings of the 2006 IEEE Congress on Evolutionary Computations (CEC 2006), July,pp. 1784-1791, 2006.
38. A Ahmadi, F Karray, and M Kamel, "Multiple cooperating swarms for data clustering", Proceedings of the 2007 IEEE Swarm Intelligence Symposium (SIS 2007), April, pp. 206-212, 2007.
39. M Slaney, M Casey, "Locality-sensitive hashing for finding nearest neighbors", IEEE Signal Processing Magazine, Vol. 25 No. 2, pp. 128-131, 2008.
40. Y Lu, S Wang,, S Li, and C Zhou, "Particle swarm optimizer for variable weighting inclustering high-dimensional data", Machine Learning, Vol. 82 No. 1, pp. 43-70, 2011.