WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Risk Factor Analysis for Health Prediction

Jessica Priya, Md Nadeem Saleem Jamadar, Mohd Huzaifa Hussian, M M Saqib Roshan

UG Students, Department of Information Science and Engineering,Guru Nanak Dev Engineering College Bidar,Karnataka
Associate Professor, Department of Information Science and Engineering,Guru Nanak Dev Engineering College Bidar,Karnataka

A B S T R A C T

In this paper, we present the techniques and applications of data mining in medicinal and instructive parts of Clinical Predictions. In medicinal and health care fields, a huge quantity of information is turning into accessible due to availability of computers. Such an oversized amount of information can't be processed to make health predictions in the early stage and make treatment schedules to diagnose. Our aim is to assess the techniques of data processing in the fields of clinical and health care to develop correct choices. It also offers a close exchange of medicinal information handling strategies which may improve various parts of Clinical Predictions. It's a latest powerful technology that is of high interest in the computer world. It uses already existing information in several databases to rework it into new researches and results. From huge data sets, to extract new patterns and the knowledge related to these patterns data mining uses machine learning and database management. Particularly the task is to get data by the means of automatic or semi-automatic. The various parameters enclosed in data processing include clustering, forecasting, path analysis and predictive analysis

Keywords: Machine Learning, Data Mining, clinical predictions, clustering, predictive analysis, forecasting

1. Introduction

Machine Learning and Data mining is a process of knowledge discovery from unknown or useless datasets. There are various techniques of data mining that are used to process the data and convert them as useful information. The data mining can be used in the various fields such as business analysis, healthcare, stock management etc. Medical field has wide amount of data that can be processed by the help of data mining techniques. It might have happened before that yourself or someone near you want immediate help of doctor but could not find anyone. By creating a model that can predict the diseases based on user symptoms is quite helpful in getting fast and appropriate medical facilities for patients. The timely analysis of data and gaining accurate prediction of diseases from symptoms can save many lives. Early detection of diseases helps doctor to give accurate medication. In the field of medicine different algorithms of machine learning are used for predicting different diseases and helps the physicians to diagnose fast. Based on the input of data the accuracy of results may vary.

There are currently a lot of health institutions that have been developed such as hospitals and medical centers which are crucial to maintain and improve the health of the community around us. It is a prime establishment of giving proper health care especially for every one of us who have ever lived. For every illness and diseases that people may face today and sometime in the future, it is because of these medical institutions and all the doctors who worked at these places that have made our lives physically better and healthy. Although hospitals now are well-equipped with their staffs working, there are still known issues that persist that cause the staffs to make the poor clinical decision that affects a patient's health such as the lack of qualified doctors, unorganized health information and poor communications between doctors and patients

Here we use some intelligent data mining techniques to guess the most accurate illness that could be associated with patient's symptoms. If the system is not able to provide suitable results, it informs the user about the type of disease or disorder it feels user's symptoms are associated with. If users symptoms do not exactly match any disease in our database, is shows the diseases user could probably have judging by his/her symptoms. It also consists of doctor address, contacts along with Feedback and administrator dashboard for system operations.

1.1. Existing System

Health care institutions are essential as it provides to every single people in the world proper health care. Its main purpose is to improve the current health of the community that we have shared and created. A health care institution such as hospitals or medical centers would essentially consist of numerous of doctors that were qualified and specialize on treating patients of their current illness that they endured and to restore them to proper health. Throughout this day and age, new technologies have been created and developed to improve people's daily life and routine, especially for health care. Doctors and nurses were now guided by smart health prediction system on the purpose of storing medical information that may be used for research and diagnosis. A few years ago, doctors were expected to use their intuition and experience to handle every medical situation that different patients are facing every day. Although their current approach may have saved people's lives back then, they are still prone to errors and wrongdoings that have endangered human life4 . It is without a doubt a heavy burden for everyone especially the medical staffs to understand that several decisions could heavily affect other people's lives and health, it is also why such system itself proves to be vital on guiding medical staff to make a proper clinical decision to cure and restore the human health. A smart health prediction system is defined as a healthcare system that is intended to assist health professionals in their decision-making process regarding medical situations. This system will provide the guidance and information needed for doctors to diagnose patient on their medical illness and it will eliminate the difficulties that the doctors needed to encounter, particularly in their clinical decision-making process. The system would require to gather a whole lot of medical information that are valuable to be used on predicting a patient's health status, these patterns of information will be analyzed by using data mining techniques in order to find correlations and discover new pieces of information from unstructured data. By using data mining tools, it will not only be able to produce reliable results with less time consumption and complexity but also with smart decision-making and useful information..

1.2. Proposed System

To beat the downside of existing framework we have created smart health prediction System. We have built up a specialist framework called Smart Health Prediction framework, which is utilized for improving the task of specialists. A framework checks a patient at initial level and proposes the possible diseases. It begins with getting some information about manifestations to the patient, in the event that the framework can distinguish the fitting sickness, at that point it proposes a specialist accessible to the patient in the closest conceivable territory. On the off chance that the framework isn't sufficiently sure, it asks few questions to the patients, still on the off chance that the framework isn't sure; at that point it will show a few tests to the patient. In light of accessible total data, the framework will demonstrate the result. Here we utilize some intelligent mining methods to figure the most precise disorder that could be associated with patient's appearances and dependent on the database of a couple of patients restorative record, calculation (Naïve Bayes) is connected for mapping the side effects with conceivable diseases.This framework improves undertaking of the specialists as well as helps the patients by giving vital help at a soonest organize conceivable

Here we use some intelligent data mining techniques to guess the most accurate illness that could be associated with patient's symptoms. If the system is not able to provide suitable results, it informs the user about the type of disease or disorder it feels user's symptoms are associated with. If users symptoms do not exactly match any disease in our database, is shows the diseases user could probably have judging by his/her symptoms. It also consists of doctor address, contacts along with Feedback and administrator dashboard for system operations.

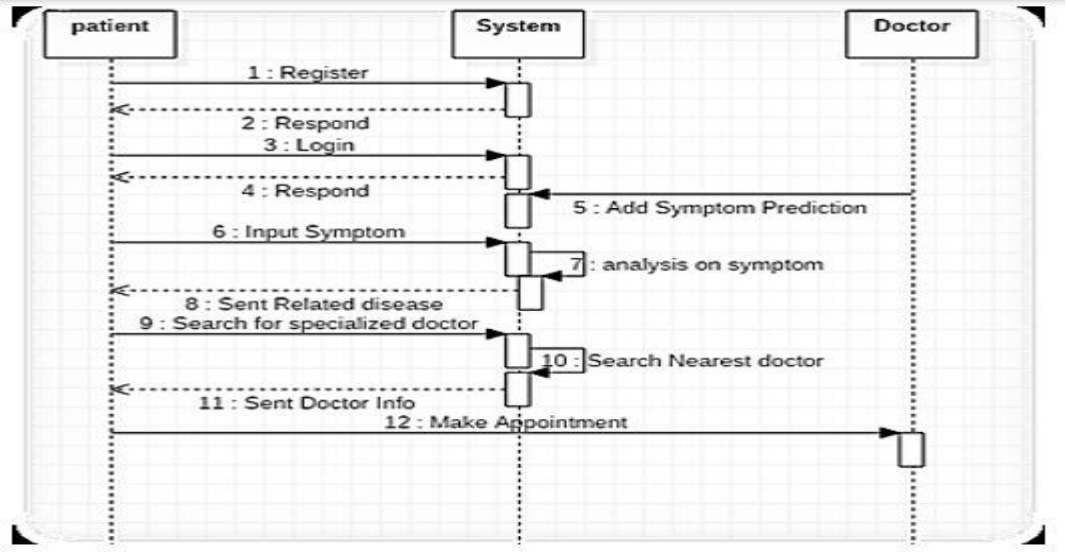

Fig 1.2.a:Sequence diagram for health prediction system

2. Data Mining Architecture

Data mining is a process of discovering analyzing different data patterns from large raw datasets. The main aim of data mining is to extract the relevant information from comprehensive dataset. The data mining comes with a bundle of packages such as machine learning, statistics and database system. All this factors determine the efficiency in Knowledge Discovery in database process. KDD consist of various process such as data cleaning, data selection, data integration, data transformation, data pattern searching and finally knowledge representation. The data mining techniques that mainly used are Association rule, Clustering, Classification, regression etc.

- The association rule can be used to establish relationship between two variables.
- The clustering is a process of grouping the structures based on similarity between them.
- The classification is assigning items in collection to target datasets.
- The regression tries to estimate the various mode to find the relation between data with least error.

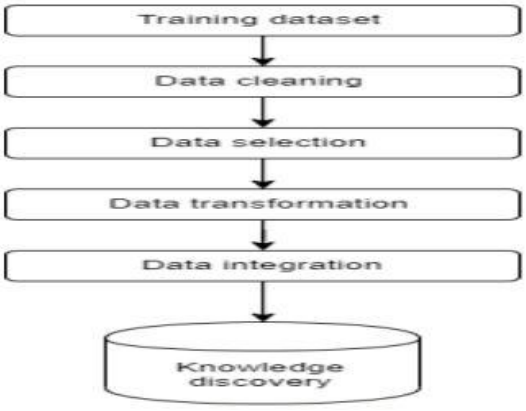

Fig 1.3.a:Prediction Process

2.1. Naive Bayes Algorithm

The proposed framework utilizes a data mining strategy "Naive Bayes classifier" for the development of the expectation framework. This framework includes a higher number of data indexes and characteristics which are legitimately gathered from specialist's data for the exact expectation of the symptom."Naive Bayes or Bayes" Rule is the reason for some, AI and information mining strategies. The standard is utilized to make models with prescient capacities. It gains from the "proof" by figuring the connection between's the objective (i.e., subordinate) and other factors.

Naïve Bayes Algorithm:

Following advances are actualized in Bayes calculation:-

Bayes' Theorem : $P(c \mid x) = P(x \mid c)\, P(c) / P(x)$

Where,

$P(c \mid x)$ = Posterior Probability

$P(c)$ = Prior probability

$P(x \mid c)$ = probability of predictor

$P(x)$ = Predictor's prior probability

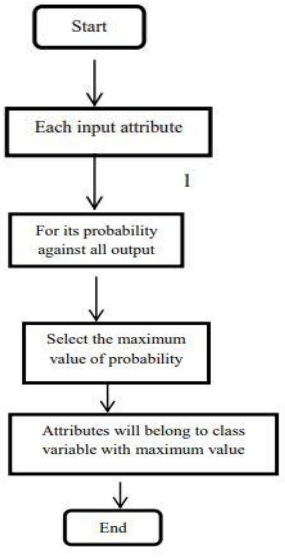

Fig 3: Naïve Bayes Flow Chart

3. Result and Snapshots

- Patient registration: If a patient is a new client, the framework asks for personal details by giving client ID and secret key through which he can login to the framework
- Prediction of the disease: The patient will show the side effects caused as a result of his ailment. The system will make certain request with respect to his ailment and after that anticipate the disease depending on the indications determined by patient and the framework will likewise propose specialists dependent on the illness.
- Doctor Consultation: Patient can scan for specialist by indicating name, address or type.and consult immediately
- Doctors feedback:Input: Patient will comment his view and this will be accounted to the administrator

Predictions–

4.Conclusion

Data mining can be helpful in the field of restorative space. Anyway protection, security and unfit to sign into the record are the huge issues on the off chance that they are not tended to and settled appropriately. It portrays the proposition of a crossover information mining model to separate arrangement learning for the guide of different maladies in the clinical choice framework and presents a structure of the apparatus different devices utilized for investigation. Now and again the circumstance happens when you need the specialist's assistance promptly, however they are not accessible because of some reason. In our venture, we have planned another well being forecast framework, which is an online framework, and different patients from any areas can see it. Our framework involves fundamental parts, for example, quiet login, enter side effects in the System, and recommend medications, proposes an adjacent specialist. The application takes the contribution of different manifestations from the patient, does the examination of the entered side effects, and gives fitting sickness expectation. Our framework enables the clients to get an examination of the indications they give for anticipating the malady they are experiencing. Some of the time the circumstance happens when you need the specialist's assistance quickly, yet they are not accessible because of some reason. Along these lines, it enables the clients to get an examination of the side effects they give for anticipating the infection they are experiencing.

5 Future Scope

Concealed learning will be extracted from the verifiable information in the proposed framework, by getting ready datasets by applying apriori calculation. Anticipating savvy wellbeing should be possible just if framework reacts that way. These datasets will be contrasted and the approaching questions and the last report will be produced utilizing Association Rule Mining. Since this proposed system will chip away at genuine chronicled information, it will give exact and productive outcomes, which will enable patients, to get the conclusion in a split second. More work should be possible later on by utilizing more informational index identified with heart sicknesses and by utilizing diverse information decrease techniques to improve the characterization. For better precision and expectation of heart sicknesses the datasets that will be used must be quality organized and free from special cases, inconsistencies, and missing characteristics.

This web application can be additionally upgraded in an Android application. This will be accessible to clients on versatile premise and its utilization can be additionally expanded. Likewise, highlight like getting the specialist online on a visit with the goal that patients can straightforwardly converse with the concerned specialists. The modules doing malignant growth examination can be coordinated to discover how close the individual related with disease is. This will make this web application unsurprising in obvious sense.

REFERENCES

1. Palaniappan S et al. Dept. of IT, Heart disease prediction usingtechniques of data mining, IEEE .
2. “Survey on Data Mining Algorithms in Disease Prediction”,Kirubha Vetal. IJCTT, 2016
3. “Performance of data mining techniques”,Abdelkrim Haqiq et al.International journal of Database Management Systems (IJDMS), june2016.
4. “ Survey on data mining techniques for diagnosis of diseases in medical domain”,R siddarth, I parvathi international journal of computer science and information technologies, 2014.
5. Narendra Ku. J lambodar, distributed data mining classification algorithms for CKD prediction, International journal of emerging research in management & technology
6. “Using machine learning algorithms in cardiovascular disease risk evaluation”,Sitar-Taut, V.A., etal., Journal of Applied Computer Science & Mathematics, 2009
7. Ghosh Joydepp, S Michael et al. Top ten algorithms in data mining,2007
8. G Attardi et al., Automatic web page categorisation by link and content analysis,1999.
9. “Data Mining and Data Warehousing”,Ho, T. J. Prentice Hall, 2005.
10. “Enhancing Naive Bayes with Various Smoothing Methods for Short Text Classification”,Cong,Nadia et al. ACM, 2012
11. W. Aiello, S. Lodha, and R. Ostrovsky, “Fast digital identityrevocation,” in Advances in Cryptology–CRYPTO 1998. Springer,1998, pp. 137–152.
12. D. Naor, M. Naor, and J. Lotspiech, “Revocation and tracing schemes for stateless receivers,” in Advances in Cryptology– CRYPTO 2001.Springer, 2001, pp. 41–62.
13. C. Gentry, “Certificate-based encryption and the certificate revocation problem,” in Advances in Cryptology–EUROCRYPT 2003. Springer, 2003, pp. 272–293.
14. V. Goyal, “Certificate revocation using fine grained certificate space partitioning,” in Financial Cryptography and Data Security. Springer, 2007, pp. 247–259.
15. A. Boldyreva, V. Goyal, and V. Kumar, “Identity-based encryption with efficient revocation,” in Proceedings of the 15th ACM
12. Sinha, Parul, and Poonam Sinha. "Comparative study of chronic kidney disease prediction using KNN and SVM." International Journal of Engineering Research and Technology, Vol 4, no. 12 pp. 608-12, 2015.
13. Kumar, Manish. "Prediction of chronic kidney disease using random forest machine learning algorithm." International Journal of Computer Science and Mobile Computing, Vol 5, no. 2, pp. 24-33, 2016