INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744
VOL 4 ISSUE 3 (2018) PAGES 3525 – 3533
RECEIVED : 05.04.2018 PUBLISHED : 08.05.2018
May 08, 2018

EHR STORED ON BLOCKCHAIN USED TO PREDICT HEART RELATED DISEASES

[1]Nandish R, [2]Minith D Jain, [3]Manjunath C R, [4]Sahana Shetty
[1,2]Department of Computer Science And Engineering, School Of Engineering And Technology, Jain University, India
[3]Associate Professor, Department of Computer Science And Engineering,School Of Engineering And Technology, Jain University, India
[4]Assistant Professor, Department of Computer Science And Engineering,School Of Engineering And Technology, Jain University, India

ABSTRACT

Heart related diseases are one of the leading causes of death which can be prevented if found at an early stage and treated accordingly. The predictive analysis is significantly useful in detecting diseases at an early stage using electronic health records, based on various kinds of factors like style of living, genealogy, addiction to smoking and health conditions like BP and diabetes. The amount of patient data produced by healthcare industry is massive. Therefore, there is need for data processing for mining these information sets to supply recommendation to patients. An Electronic Health Record system is crucial to produce insights and anticipate the results from already existing patient data. EHR system is useful in the processing of immense quantity of the information, in order to predict disease based on the results obtained. One of the major problems with EHR data is that it is scattered across various healthcare systems, which can put the patient's data at risk and also, would make sharing of records all the more difficult. Therefore, we suggest the use of Blockchain to store the Electronic health records, which will be useful in accessing patient data which is scattered across various healthcare system and hence, can be used to predict diseases based on various algorithms

INTRODUCTION

Big data analytics can be defined as a technique of comparing massive and varied amount of information or data sets to uncover hidden patterns, unknown relationships, market drift, client choices and alternative useful information that may aid organizations to build highly-informed business selections. The purpose of analysing the huge amount of data that is a part of the healthcare is such that, one can generate better health reports and highly precise predictive models around individual patients so that we can better diagnose and treat disease.

An Electronic health record (EHR) is a electronic version of a patient's health report which is maintained over time either by the doctor or the patient himself. EHRs are patient-centric records that make patient information available instantly and securely to any authorized individual. While an EHR does contain the medical and treatment histories of patients, an EHR system is built to go beyond standard clinical data collected in a provider's office and can be inclusive of a broader view of a patient's care. EHRs can:

· Contain a patient's medical history, diagnoses, medications, treatment plans, immunization dates, allergies, radiology images, and laboratory and test results
· Allow access to evidence-based tools that providers can use to make decisions about a patient's care
· Automate and streamline provider workflow

Blockchain is a gradually growing list of records or blocks, which are connected to each other using cryptography. In a blockchain, each block is connected to the previous using a cryptographic hash value and possibly a timestamp variable, this is done to make sure that the content of block cannot be changed without altering the other blocks. Blockchain works on the principle of distributed ledgers which is handled by peer to peer network. Blockchain provides more data transparency and trust.

Typically, blockchain were designed as crypto-currency, which solved the "double spending problem" of the digital currencies. But the concept of blockchain can be used in various fields such healthcare, educational systems.

Blockchain in healthcare, the major issues with the electronic health records is scalability, privacy of the records and access permission. These issues can be solved using the concept of blockchain which provides more scalable and secure way to store user records which can be accessible from anywhere.

Heart disease includes variety of circumstances that could affect the working of the heart. Diseases which are related to the heart are blood vessel diseases like the coronary artery disease, heart rhythm complications, and heart related defects that you're born with, and many others.

Factors that influence heart related issues/diseases:

- Smoking: Smoking is one of the controllable factors that cause heart related diseases. Smoking increases the chances of getting heart related diseases through many mechanisms. Smoking causes blood clotting, it damages the blood vessels. Nicotine, which is the main component of a cigarette, leads to increased heart rate and increases the blood pressure.
- Being overweight or obese: Narrowing of the arteries(atherosclerosis), hypertension and diabetes can be caused as a result of being overweight. Being overweight is one of the major reasons that influence heart related diseases.
- Being physically inactive: Physical inactivity is one of the major factor that influences heart related diseases. Studies have shown that being physically active reduces the risk of heart related diseases by around 30%. Being physically active keeps your weight regulated and improves blood circulation in the body which reduces the chance of having heart related disease. Smoking also increases the risk of stroke and impotence. Quitting smoking can effectively reduce the chance of getting heart disease.
- High blood pressure or Hypertension: One of the major risk factor for stroke is Hypertension. Hypertension causes the blood vessels to clog or weaken which could lead to a potential heart attack. Diabetes causes damage to nerves as well as blood vessel and there is no indication such as chest pain when a diabetic person gets a heart attack.

- High blood cholesterol: High level of LDL cholesterol can lead to heart attacks and ischemic strokes. Whereas HDL reduces the risk of a stroke since it carries cholesterol away from the bloodstream.
- Diabetes and prediabetes: Diabetes is one of the factors causing cardiovascular diseases that cannot be modified through human effort. A person having diabetes is more likely to develop heart related diseases than people without diabetes.
- Having a family history of early heart disease: A person has higher risk of developing heart related disease if both parents have had heart diseases before the age of 55. Also, one of the factors that causes heart related diseases that is cholesterol level, is inherited from parents, increasing the chance of developing heart diseases.
- Unhealthy diet: A healthy diet is one of the important factor to reduce the risk of heart related disease. A diet containing high saturated fats increase the risk of developing CVD.

Some of the different types of Heart Diseases

Congenital heart defect

This is a heart problem which effects the walls, valves and vessels of the heart and is generally present since birth. Symptoms of the disease may either appear during the early years shortly after birth or may it not appear until many years after birth.

Symptoms include:

- Fainting, swelling, breathing problems, dizziness, varying heart beats
- Chest pains, low birth weight, delayed growth, breathing problems in the case of the newborns

Arrhythmia

Arrhythmia is a heart disease which is described by irregular heartbeat. Arrhythmias is the condition when the heartbeat is irregular caused due to improper working of the electrical impulses that coordinate the heartbeat. This causes the heart to beat in a fashion it shouldn't be beating, that may be too fast, too slowly. There are several ways in which a heartbeat can vary. These are as follows:

- When the heartbeat is very fast i.e the heart beat rate is more than 100 beats per minute. This is known as Tachycardia
- When the heart beats is very slow i.e the heart beat rate is less than 60 beats per minute. This is known as Bradycardia
- abnormal beats caused due to problems at ion level in the heart
- When the heartbeat is irregular, it is known as Fibrillation. This can either affect the ventricle or the atrium

The factors that are the root cause of this disease are:

- Alcoholism
- Drug abuse
- High blood pressure
- Diabetes

Myocardial infarction

Which is more commonly known as a Heart Attack. A Heart attack occurs due to interrupted blood flow or stopping of blood flow that damages or destroys part of the heart muscle. This is usually the result of a blood clot that is developed in one of the coronary arteries and it can also occur if an artery suddenly narrows or spasms. The major reason that can lead to heart attacks are improper diet, age, less physical activity or it can also be genetic.

Congestive heart failure

Which are commonly known as heart failure, heart failure is caused when the heart does not pump blood around the body efficiently and there is lack of blood needed by the organs to function properly. The reason for heart failure can be due to coronary artery disease or high blood pressure,which over time, leaves the heart too stiff or weak to fill and pump properly.

PROBLEM DEFINITION

Heart is one of the most important organ of the human body which is responsible for pumping blood through the blood vessels of the circulatory system. Some of the symptoms like high blood pressure, Diabetes, stress can show early signs of heart related diseases. Using the data from the electronic records which is stored with the help of the blockchain technology and few of the data analysis techniques, heart related diseases can be predicted at an early stage and hence, steps can be taken to prevent/cure the disease.

Implementation of EHR systems using blockchain in medical field

Currently, Electronic Health Record (EHR) systems have been broadly used in telehealth fields with the development of the Web-based and Mobile technologies. There are a variety of advantages with implementing EHR systems using blockchain, such as improving the efficiency of the diagnosis processes, achieving master data management, real-time information management, more privacy of data being added, easier access to the information and higher scalability. From a perspective of physicians, adopting EHR can enable them to obtain patients' medical records that can be used as a reference for diagnoses. The medical information stored in the database is expected to play a positive role in medical decision making. Using EHR systems implemented using Blockchain would benefit patient as well. They would gain control over the documents that are being shared. Usually Hospitals or Doctors are in control of the electronic health records. But Blockchain would provide control of the records to the patients.

EXISTING SYSTEM

In Big data, computational techniques and sources are swiftly increasing and few of the uses for well-being and bettering health are multifarious. The main aim for healthcare is to give out a continuously learning infrastructure to develop a system which is participatory, preventive and predictive along with real-time knowledge production. The tremendous potential of Big data analysis has ability to revamp healthcare and has ability to change the populations health. By utilizing this potential successfully, ability of it depends on solving the associated challenges with data security, privacy, governance and ownership.

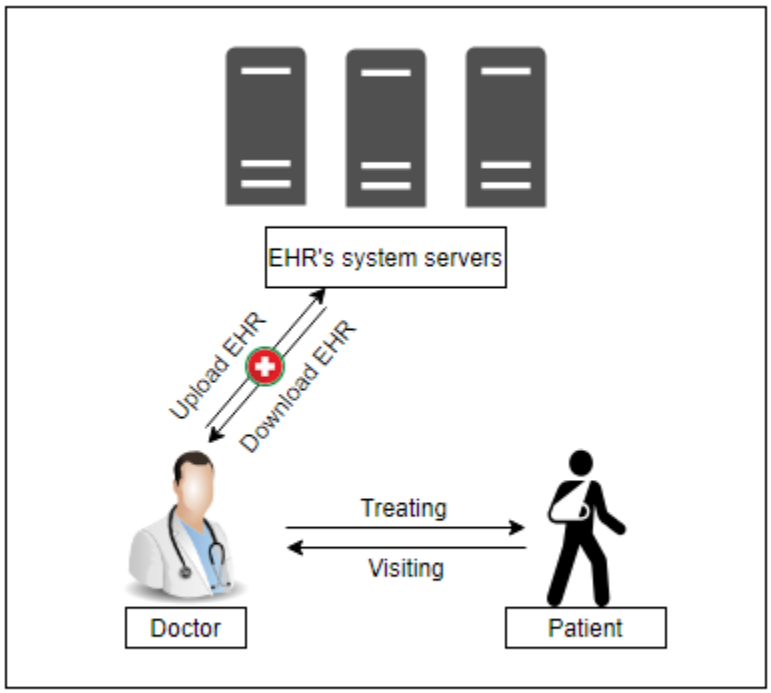
Most of them focus on problem of preventing diagnosis mistakes in digital health record systems and they have came up with a solution which is able to overcome diagnosing medical errors in big data. The mechanism they came up with was on OEHR-EPM which was created in order to achieve the predicted prevention goal with the help of an ontological approach. In the proposed model the major algorithm used was Error Prevention Adjustment Algorithm (EPAA) which was a paperback schematic representation done using ontology. The proposed model efficiency was proved using experimental evaluations. They are also trying to address research limitations, such as to improve calculation and possible messy statistics.
A model for prediction have been used J48 decision tree to classify heart disease based on the clinical attributes. The precision of Pruned J48 Decision Tree with REP approach is much better than the simple Pruned and Unpruned approach. Output obtained shows that fasting blood sugar is the most consequential attribute which gives better classification against the other attributes but it does not give better accuracy.

One of the existing system suggests instructions for the less amount data and data type required by using longitudinal electronic health records data to teach effectual disease predictive models.

The clinical as well as treating decisions in healthcare is extended by predictive analysis which is the main/key technique. Major challenge is accessing the data from very large database, which is scalable to petabytes, to house data of the patients obtained from different origins. Techniques like Hadoop and Hive are used to scale the data, since the data obtained from the database are very large data must be scaled, At this point Hadoop and Hive comes into picture. Hadoop uses programming model called MapReduce which is used for processing Huge/Large sets of data. Data in Hive is stored in a distributed system, SQL queries are used to retrieve the data. Hives architecture is scalable for multiprocessors remarkably huge than existing ones.Architecture is promising that it helps in obtaining significant data from patients. If the patients with the high-risk are identified, they can be provided with individualised medicines and also to give inputs to the lifestyle of the patient to reduce the risk.

All the Existing systems just include either adopted Electronic health records for all the diseases which contains high dimensional data containing massive information and big data analytics can extract such knowledge to improve the quality of health care or they have just used few of the machine learning algorithms to predict future detection of heart failing using health records stored digitally and also heart disease prediction using predictive modelling techniques like Decision tree.

INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744
VOL 4 ISSUE 3 (2018) PAGES 3525 – 3533
RECEIVED : 05.04.2018 PUBLISHED : 08.05.2018
May 08, 2018



The Present EHR System

PROPOSED SYSTEM

Earlier, the doctors were completely depended on the patient's medical history and on the physical examinations done on the patients. The quality of the treatment was determined by each doctor's ability to get the information and process it accordingly. In this modern era, the quantity of objective information originating from complex diagnostic imaging tools, advanced laboratory results has increased exponentially, and also, the sources of obtaining this data has changed, such as from wearable devices, mobile devices, and various other technologies. And over the recent years, electronic health records (EHRs) have assisted doctors in utilising the desired data more efficiently and precisely.

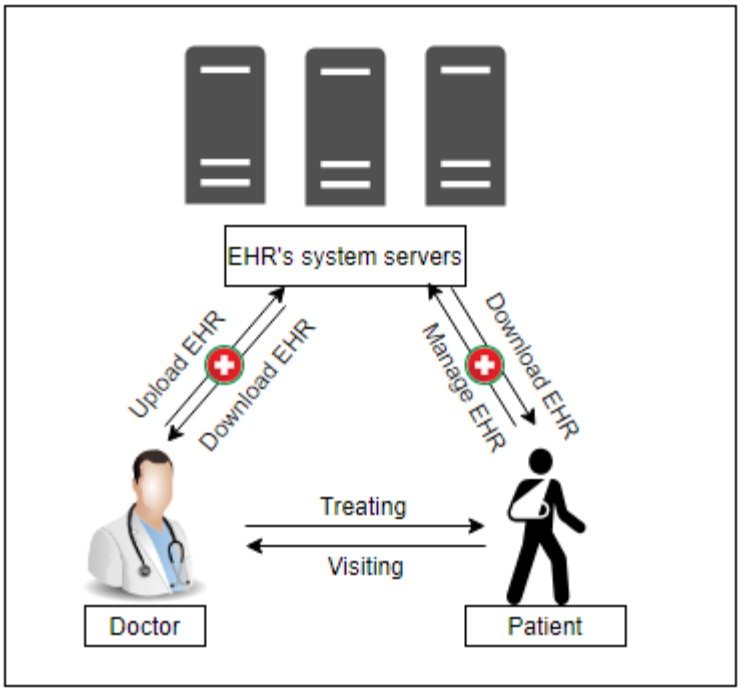
In this paper, we suggest the use of blockchain technology to store the electronic health records in order to provide scalable, easier to access and more private records. Health care records usually contain reports, images, documents and various other stuff, so storing these records across various health care systems would be a waste of storage space. So, we suggest the use of blockchain to store all the records in single data pool, where all the patients would store their health records and the doctors can access the records of a particular patient with the permission of the patient. The data pool could support text mining, interactive queries and machine learning which could be used in predicting diseases using data analytic techniques. Using of blockchain would ensure that the records cannot be accessed by anyone who doesn't have the permission to access the records. This could be implemented using public and private key encryptions. Whenever a record is created or a content of a record is changed, it is stored in blockchain with the user's unique identifier and a timestamp. So that the records updated cannot be reverted back.

INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744
VOL 4 ISSUE 3 (2018) PAGES 3525 – 3533
RECEIVED : 05.04.2018 PUBLISHED : 08.05.2018
May 08, 2018

In terms of data privacy and access security, the user will have complete control over what data is being shared and who is accessing the said data. The user could also set permissions for the data that is to be accessed like the content that is being accessed, time frame for accessing the data.

Proper use of big data in medical care possesses great promise and hence, it can be tested on various stages of research in the field of cardiology, like health management of large-scale population, risk prediction of heart diseases, precision medicine, and clinical decision support through machine-learning algorithms.

One of the models that can be used for prediction of the heart related diseases is the "decision tree model", where analysis on a predictive characteristic is denoted by the internal node and the attribute values are represented by a branch in the tree. Predicted classes or class distributions are denoted by a leaf node. An unlabelled object is defined by starting at the root node of the tree, then traversing through the tree structure, based on the values of the predictive attributes in this object. In the decision tree algorithm, the assumption is that the data objects are described by a set of characteristics that are fixed, where every predictive characteristic considers a small quantity of disjoint values and the target variable has discrete output values, where class label is represented by each value.



The Future EHR System

INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744
VOL 4 ISSUE 3 (2018) PAGES 3525 – 3533
RECEIVED : 05.04.2018 PUBLISHED : 08.05.2018
May 08, 2018

METHODS USED FOR DIMENSIONALITY REDUCTION OF EHR DATA:

TABLE I

Method	Advantages	Limitations
Feature extraction: PCA, SVD, tensor-based approaches*	Reduces dimensionality; relatively immune to noise	Performance usually inferior to supervised approaches; difficult to interpret results
Feature selection: filter based (mRMR), wrapper based (sequential feature selection) *	Reduces dimensionality; easy to interpret	Sometimes affected by noisy data

Methods for Predicting heart related diseases using Big Data Analytics

(i) Big Data Analytics for heart related diseases.
Majority of the health Records that are stored digitally contains high dimensional data which not only requires longer computing time but also affects the precision of the analysis. Hence, we try to reduce dimension of the data by exploring the subset of the variables that contains as much of the features of the original data as viable with two feature methods:
(i) Feature selection focuses on selecting optimal subset of existing features.
(ii) Feature extraction focuses on transforming prevail features to a more concise set of dimensions.

(ii) Pre-processing of EHR Data stored digitally.
Information which is embedded in EHR is massive but not organized. Hence, digital health record data needs data to be preprocessed systematically. For data missing on EHR, standard methods either chalk up the values that are missing by median or mean in a population. These methods are easy and simple for implementing them, but they disregard the structure of the data and has a tendency of introducing additional biases. Hence, more vigorous data which is missing, its imputation methods are needed.

REFERENCES

[1]. Predictive Analytics on Electronic Health Records (EHRs) using Hadoop and Hive (2015) - Haritha Chennamsetty, Suresh Chalasani, Derek Riley

[2]. Omic and Electronic Health Records Big Data Analytics for Precision Medicine (2016) - Po-Yen Wu, Chih-Wen Cheng, Chanchala D. Kaddi, Janani Venugopalan, Ryan Hoffman, and May D. Wang
[3]. Electronic Health Record Error Prevention Approach Using Ontology in Big Data (2015) - Keke Gai, Meikang Qiu, Li-Chiou Chen Meiqin Liu
[4].K. Srinivas, B.Kavihta Rani, A.Govrdhan, Applications of Data Mining Techniques in Healthcare and Prediction of Heart Attacks, (UCSE) international Journal on Computer Science and Engineering Vol. 02, No. 02,250-255,2015.
[5].Anbarasi M, Anupriya E, Iyengar NCHSN. Enhanced prediction of heart Disease with feature subset selection using genetic algorithm. International Journal of Engineering Science and Technology;2:5370-76,2015.
[6].Kenney Ng, PhD; Steven R. Steinhubl, MD; Christopher deFilippi, MD; Sanjoy Dey, PhD; Walter F. Stewart, PhD, MPH. Early Detection of Heart Failure Using Electronic Health Records 2016;9:00-00,2015.
[7].Laure A. Linn; Martha B. Koo, M.D.Blockchain For Health Data and Its Potential Use in Health IT and Health Care Related Research,2017
[8]Xiao Yue; Huiju Wang; Dawei Jin; Mingqiang Li; Wei Jiang.Healthcare Data Gateways: Found Healthcare Intelligence on Blockchain with Novel Privacy Risk Control,2016
[9].Asaph Azaria; Ariel Ekblaw; Thiago Vieira; Andrew Lippman. MedRec: Using Blockchain for Medical Data Access and Permission Management,2016
[10].Rui Guo; Huixian Shi; Qinglan Zhao; Dong Zheng.Secure Attribute-Based Signature Scheme with Multiple Authorities for Blockchain in Electronic Health Records Systems,2017
[11].Kenney Ng, PhD; Steven R. Steinhubl, MD; Christopher deFilippi, MD; Sanjoy Dey, PhD;
Walter F. Stewart, PhD, MPH. Early Detection of Heart Failure Using Electronic Health Records,2016
[12].Sana Bharti; Dr.Shailendra Narayan Singh; Analytical Study of Heart Disease Prediction Comparing With Different Algorithms,2015
[13].Joonseok Kim, MD. Big Data, Health Informatics, and the Future of Cardiovascular Medicine,2017
[14].Ankita Dewan; Meghna Sharma. Prediction of Heart Disease Using a Hybrid Technique in Data Mining Classification,2015
[15]Monika Gandhi; Dr.Shailendra Narayan Singh. Predictions in Heart Disease Using Techniques of Data Mining,2015