Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 6.017

IJCSMC, Vol. 6, Issue. 4, April 2017, pg.168 – 172

HEART DISEASE PREDICTION USING ANN ALGORITHM IN DATA MINING

Mr. P Sai Chandrasekhar Reddy[1], Mr. Puneet Palagi[2], Ms. Jaya[3]

[1]M. Tech Student, Department of Computer Science and Engineering, New Horizon College of Engineering, Bangalore, India
[2]Assistant Professor, Department of Computer Science and Engineering, New Horizon College of Engineering, Bangalore, India
[3]Sr. Assistant Professor, Department of Computer Science and Engineering, New Horizon College of Engineering, Bangalore, India
[1] pulisaichandrasekharreddy@gmail.com; [2] puneetpalagi@gmail.com; [3] jayamanojkumar@gmail.com

---

Abstract:- Data mining along with soft computing techniques helps to unravel hidden relationships and diagnose diseases efficiently even with uncertainties and inaccuracies and it was using algorithms in that data mining mainly Coronary Heart Disease. It is an automatic process which scans the infected body part of human system and find out reason behind the sickness and disease. Heart disease is the leading Causes of death in the world over the past ten years. It is must that diagnosing system must generate accurate results so that proper treatment can be available for the patient. In earlier systems various techniques was used for predicting the reason behind the disease which lacks in accuracy of the output. Various heart diseases like heart attacks, chest pain etc. This thesis is based on the heart disease diagnosis of patients. Heart disease is a prevailing disease nowadays. Now due to increasing expenses of heart disease, there was a need to develop a new system which can predict heart diseases in an easy and cheaper way. The proposed work divides proposed system in two parts such as performance model and prediction model. Performance model is designed to evaluate the overall performance of the application. Prediction Model is used to predict the condition of the patient after evaluation on the basis of various parameters like heart beat rate, blood pressure, cholesterol etc. The accuracy of the system is proved in java.

Keywords: data mining, Artificial Neural Network, PSO.

I. INTRODUCTION

A major challenge facing healthcare organizations and industries (hospitals, medical testing centre) is the provision of quality services at affordable costs and Quality service implies diagnosing patients correctly and administering treatments that are effective. In the Poor clinical decisions can lead to disastrous consequences which are therefore unacceptable. Hospitals must also minimize the cost of clinical tests and reports. They can achieve these results by employing appropriate computer-based information and/or decision support systems. . Diagnosing is a process which is done by automated systems or machines in human system to find out the nature of the disease by monitoring the various symptoms of the illness. Diagnosing is a most complicated task to perform. Diagnosing machines or systems are quiet helpful in this process because not every doctor must have the knowledge of each and every kind of problem of disease.

In this diagnosing manually by doctors can lead to inaccurate results some times. Thus an automated diagnosing machine is used by them to diagnose the problem accurately. The WHO consortium has shared this information that ten millions of deaths occur in this world is just because of heart disease.so it was a very

dangers problem in world. These systems typically generate huge amounts of data which take the form of numbers, charts and images. Unfortunately, these data are rarely used to support clinical decision making.

There is a wealth of hidden data in these data that is largely untapped. This raises an essential question: "How can we turn data into suitable information that can enable healthcare consultants to make intelligent clinical decisions of process?" This is the main motivation to implement for this paper. Heart disease is the main reason behind the deaths in various countries of the world and even in India also. In case of USA one person is dying every 36 seconds just because of some kind of heart disease. Most of people died in heart disease problem only. There are many types of heart disease such as coronary heart disease, cardiovascular disease and cardiomyopathy disease. Cardiovascular disease is that disease which directly effects the blood circulation in the body and blood vessels which are connected to the heart. So we can try to decrease in this heart disease. Cardiovascular disease leads to various illnesses in the body like high BP, coronary artery disease, stroke etc. Cardiovascular disease can also leads to the death of the person also.

It is a type of coronary disease in which oxygen and blood is not properly is not supplied to the heart because of reduction in the size of coronary arteries. Coronary heart disease also includes heart attacks and chest pain. Medical diagnosis is a difficult task that requires operating accurately and efficiently. Increase in heart disease is because of many facts like great BP, smoking, Family history etc. some other factors that also causes heart diseases are high cholesterol level, hyper stiffness, improper diet etc. Symptoms of heart disease are like:

- Peripheral artery disease
- High blood pressure
- Cardiac arrest
- Congestive heart failure
- Arrhythmia
- Peripheral artery disease
- Stroke
- Fatigue
- Pain in chest
- Palpitation

II. DATAMINING ALGORITHMS

Data Mining is about explaining the past and predicting the future by means of documents analysis. Datamining is a multi-disciplinary field which combines measurements, machine learning, artificial intelligence and database technology. The value of data mining presentations is often estimated to be very high. Many businesses have stored large amounts of data over years of operation, and data mining is able to extract very valuable understanding from this data. The businesses are then able to leverage the extracted knowledge into more clients more sales, and greater profits. This is also true in the engineering and medical fields. Data mining predicts the future of modelling. . The abstraction of information is followed by several other procedures such as Data Cleaning, Data Integration, Data Conversion, Data Mining, Pattern Evaluation and Data Presentation. After all these processes are completed, we are now in a position to use this information in many applications such as Fraud Finding, Market Analysis, Production Control, Science Investigation etc.

1. ARTIFICIAL NEURAL NETWORKS:

ANN algorithm used in computer science and other research disciplines, which is based on a large collection of simple neural units (artificial neurons), loosely analogous to the observed behaviour of a biological brain's axons. Each neural unit is connected with many others, and links can enhance or inhibit the activation state of adjoining neural units. Each individual neural unit computes using outline function. There may be a threshold function or limiting function on each connection and on the unit itself, such that the signal must outshine the limit before propagating to other neurons. These systems are self-learning and trained, rather than explicitly programmed, and excel in areas where the solution or feature detection is difficult to express in a traditional computer program.

Neural networks typically consist of multiple layers or a cube design, and the signal path traverses from the first (input), to the last (output) layer of neural units. Back propagation is the use of forward stimulation to reset weights on the "front" neural units and this is sometimes done in combination with training where the correct result is known. Dynamic neural networks are the most advanced, in that they dynamically can, based on rules, form new connections and even new neural units while disabling others.

The goal of the neural network is to solve problems in the same way that the human brain would, although several neural networks are more abstract. Recent neural network projects typically work with a few thousand to a few million neural units and millions of connections, which is still several orders of enormousness less complex than the human brain and closer to the computing power of a worm. New Heart research often stimulates new patterns in neural networks. One new approach is using connections which span much further and link processing layers rather than always being localized to together neurons.

2. PARTICLE SWARM OPTIMIZATION:

PSO is a computational method that optimizes a problem by iteratively trying to improve a candidate solution with regard to a given quantity of quality. It solves a problem by having a population of candidate solutions, here dubbed particles, and moving these particles around in the search-space according to simple calculated formulae over the particle's position and velocity. Each particle's movement is influenced by its local best known position, but is also steered toward the best known positions in the search-space, which are updated as better positions are found by other particles. This is expected to move the swarm toward the best solutions.

PSO shares many similarities with evolutionary computation techniques such as Genetic Algorithms (GA). The system is initialized with a population of random solutions and searches for optima by fill in generations. However, unlike GA, PSO has no evolution operators such as crossover and mutation. In PSO, the potential solutions, called particles, fly through the difficult space by following the current optimum particles. Each particle keeps track of its coordinates in the problem space which are associated with the best solution (fitness) it has achieved so far. This value is called best. Another "best" value that is tracked by the particle swarm optimizer is the best value, obtained so far by any particle in the neighbour of the particle. This location is called best .when a particle takes all the population as its topological neighbour, the best value is a global best and is called best. The particle swarm optimization concept consists of, at each time step, changing the velocity of (accelerating) each particle toward its best and best locations. Acceleration is weighted by a random term, with separate random numbers being generated for acceleration toward best and best locations.

III. METHODOLOGY AND PROPOSED SYSTEM

The cloud server (simply the cloud), the company who provides the M-Health monitoring service (i.e., the healthcare service provider), the individual clients (simply clients), and a semi-trusted authority (TA). The company stores its encrypted monitoring data or program in the cloud server. Individual clients collect their medical data and store them in their mobile devices, which then transform the data into attribute vectors. The attribute vectors are delivered as inputs to the monitoring program in the cloud server through a mobile (or smart) device. A semi-trusted authority is responsible for distributing private keys to the individual clients and collecting the service fee from the clients according to a certain business. model such as pay-as-you-go business model. The TA can be considered as a collaborator or a management agent for a company (or several companies) and thus shares certain level of mutual interest with the company. However, the company and TA could collude to obtain private health data from client input vectors. The proper & exact diagnosis of heart disease is important so that right treatment could be given to the patient and his/her life can be saved. The diagnoses of the heart diseases were earlier done using fuzzy logics, Neural Networks etc.

3.1 Branching Program:

It is formally describe the branching programs, which include binary classification or decision trees as a special case. i am only consider the binary separating program for the ease of exposition since a private query protocol based on a general decision tree can be easily derived from our scheme. Let v be the vector of clients' attributes. To be more specific, an attribute component vi is a concatenation of an attribute index and the respective attribute value. For instance, A||KW1 might parallel to "blood pressure: 130". Those with a blood pressure lower than 130 are considered as normal, and those above this threshold are considered as high blood pressure. The first element is a set of nodes in the branching tree. The non-leaf node pi is an intermediate decision node while leaf node pi is a label node. Each decision node is a pair (ai, ti), where ai is the attribute index and ti is the threshold value with which vai is compared at this node. The same value of ai may occur in many nodes, i.e., the same attribute may be evaluated more than once. For each decision node i, L(i) is the index of the next node if vai ≤ ti; R(i) is the index of the next node if vai > ti. The label nodes are

attached with classification information. Repeat the process recursively for ph, and so on, until one of the leaf nodes is reached with decision information.

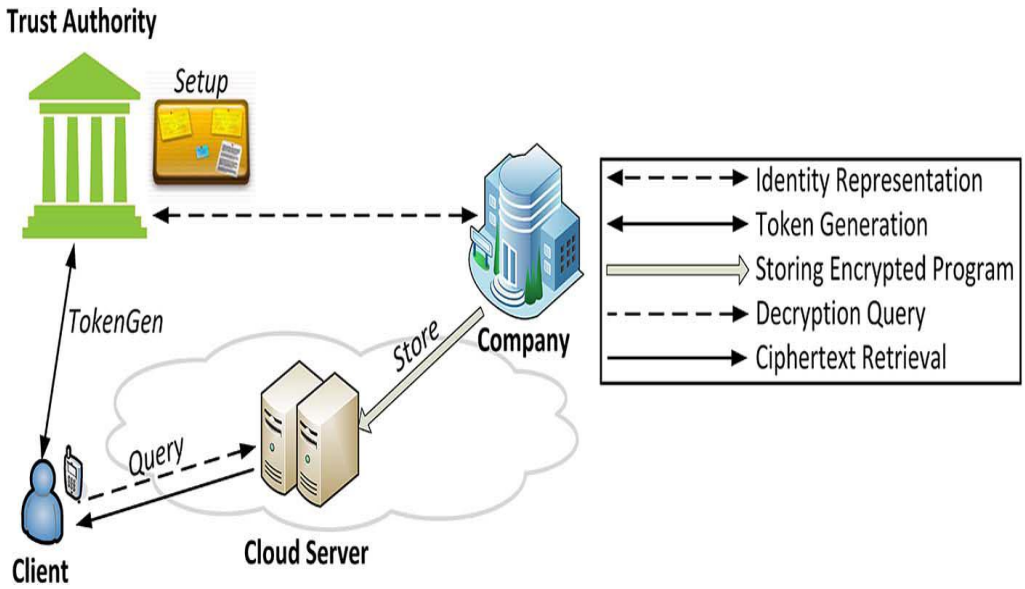
3.2 Token Generation:

To generate the private key for the attribute vector v=(v1, · · · , vn), a client first computes the identity representation set of each element in v and delivers all the n identity representation sets to TA. Then TA runs the AnonExtract(id, msk) on each identity id ∈ Svi in the identity set and delivers all the respective private keys skvi to the client.

3.3 Query:

A client delivers the private key sets obtained from the Ann algorithm to the cloud, which runs the PSO algorithm on the cipher text generated in the Store algorithm. Starting from p1, the decryption result determines which cipher text should be decrypted next. For instance, if v1 ∈ [0, t1], then the decryption result indicates the next node index L(i). The cloud will then use skv(L(i)) to decrypt the subsequent cipher text CL(i). Continue this process iteratively until it reaches a leaf node and decrypt the respective attached information.

System Architecture:



IV. CONCLUSION

Heart disease diagnosing is a difficult and important task in order to get patient's exact condition with respect to any disease or illness which is directly related to heart. Various heart diseases like heart attacks, chest pain etc. This thesis is based on the heart disease diagnosis of patients. Heart disease is a prevailing disease nowadays. Now due to increasing expenses of heart disease, there was a need to develop a new system which can predict heart diseases in an easy and cheaper way. Various methods had developed previously which had given methods to predict heart disease. The diagnosis is based on data mining processes. Data mining is a process of extraction of knowledge from the data in the database. In the database irrelevant data is present. Till now several data mining techniques namely classification, ANN algorithm and association rules are applied to the health data sets for predicting heart diseases. In current study ANN and PSO are used for the heart disease prediction. And easy to take report of using in this process.

REFERENCES

[1] P. Mohan, D. Marin, S. Sultan, and A. Deen, "Medinet: Personalizing the self-care process for patients with diabetes and cardiovascular disease using mobile telephony," in Proc. 30th Ann. Int. Conf. IEEE Engineering in Medicine and Biology Society, 2008 (EMBS 2008), 2008, pp. 755–758.

[2] A. Tsanas, M. Little, P. McSharry, and L. Ramig, "Accurate telemonitoring of parkinson's disease progression by noninvasive speech tests," IEEE Trans. Biomed. Eng., vol. 57, no. 4, pp. 884–893, Apr. 2010.

[3] G. Clifford and D. Clifton, "Wireless technology in disease management and medicine," Ann. Rev. Medicine, vol. 63, pp. 479–492, 2012.

[4] L. Ponemon Institute, Americans' Opinions on Healthcare Privacy,2015 [Online]. Available: http://tinyurl.com/4atsdlj

[5] A. V. Dhukaram, C. Baber, L. Elloumi, B.-J. van Beijnum, and P. D.Stefanis , "End-user perception towards pervasive cardiac health care services: Benefits, acceptance, adoption, risks, security, privacy and trust," in Proc. Pervasive Health, 2011, pp. 478–484.

[6] M. Delgado, "The evolution of health care it: Are current U.S. privacy policies ready for the clouds?," in Proc. SERVICES, 2011, pp. 371–378.

[7] N. Singer, "When 2 2 equals a privacy question," NewYork Times, Oct. 18, 2009 [Online]. Available: http://www.nytimes.com/2009/10/18/business/18stream.html

[8] E. B. Fernandez, "Security in data intensive computing systems,"in Handbook of Data Intensive Computing. NewYork , NY, USA: Springer, 2013, pp. 447–466.

[9] A. Narayanan and V. Shmatikov, "Myths and fallacies of personallyidentifiable information," Commun. ACM, vol. 53, no. 6, pp. 24–26,2010.

[10] P. Baldi, R. Baronio, E. D. Cristofaro, P. Gasti, and G. Tsudik, "Countering gattaca: Efficient and secure testing of fully-sequenced human genomes," in Proc. ACM Conf. Computer and Communications Security, 2011, pp. 691–702.

[11] A. Cavoukian, A. Fisher, S. Killen, and D. Hoffman, "Remote home health care technologies:Howto ensure privacy? Build it in: Privacy by design," Identity in the Information Society, vol. 3, no. 2, pp. 363–378, 2010.

[12] A. Narayanan and V. Shmatikov, "Robust de-anonymization of large sparse datasets," in Proc. IEEE Symp. Security and Privacy, 2008 (SP 2008), 2008, pp. 111–125.

[13] A. Narayanan and V. Shmatikov, "De-anonymizing social networks," in Proc. IEEE Computer Society, IEEE Symp. Security and Privacy, 2009, pp. 173–187.

[14] I. Neamatullah,M. Douglass, L. Lehman, A. Reisner,M. Villarroel,W. Long, P. Szolovits, G. Moody, R. Mark, and G. Clifford, "Automated de-identification of free-text medical records," BMC Med. Inform. Decision Making, vol. 8, no. 1, p. 32, 2008.

[15] S. Al-Fedaghi and A. Al-Azmi, "Experimentation with personal identifiable information," Intelligent Inf.Manage., vol. 4, no. 4, pp. 123–133, 2012.

[16] J. Domingo-Ferrer, "A three-dimensional conceptual framework for database privacy," Secure Data Manage., pp. 193–202, 2007.

[17] T. Lim, Nanosensors: Theory and Applications in Industry, Healthcare, and Defense. Boca Raton, FL, USA: CRC Press, 2011.

[18] X. Zhou, B. Peng, Y. Li, Y. Chen, H. Tang, and X. Wang, "To release or not to release: Evaluating information leaks in aggregate human-genome data," Computer Security-ESORICS 2011, pp. 607–627, 2011.

