Automatic Prediction of Silent Heart Attacks Using Neural Networks Based Hybrid Classification System

[1]Dr. B. Madhuravani, [2]Sri Sowmya Gudipati, [3]Ponnuru Sowjanya
[1]Department of Computer Science and Engineering, MLR Institute of Technology, Dundigal, Hyderabad, Telangana, India.
[2,3]Department of Computer Science and Engineering, School of Technology, GITAM (DEEMED TO BE UNIVERSITY), Hyderabad, Telangana, India.

ABSTRACT

Data mining method is performed on enormous databases for extraction of hidden patterns by means of machine learning techniques as a conventional and statistical analysis. Data mining in prediction of heart deceases plays a significant role in advancement of various applications like summarizing the deaths occurred by heart deceases, extracting the human health care flourishing and identifying the heart attacks occurred in world wide. The identification of heart attacks in any person is so complicated task in the medical field and the prediction takes many years. So we introduce a novel neural network based optimization technique for the prediction and identification of heart diseases in any individual person. A proposed method in existing new Convolution Neural Network based Multimodal Disease Risk Prediction (CNN-MDRP) estimation utilizes structured and unstructured data for the prediction of heart decease using culturing. Right when a bit of the data is divided then precision reduces. It will clear the defilement of information by Genetic Algorithm. The unstructured information will be changed over into the structured data with the help of Recurrent Neural Network (RNN). The proposed hybrid model obtains highest accuracy between classifiers which is much significant in medical diagnosing and minimizes energy.

Key words: Machine learning technique, data mining technique, recurrent neural networks , convolution neural network based multimodal disease risk.

1. INTRODUCTION

Heart attacks are a standout amongst those most rapid diseases of the present world. As stated by a survey, regarding more than 17.7 million passing's happen all over the globe yearly because of heart failure diseases [1]. From these deaths, present evaluation says that about 7.4 million were because of coronary diseases and 6.7 million were because of heart stroke [2]. Heart attack is a standout amongst the majority dangerous diseases, that can hit particular person down at any instance of time through without any intimation and silent heart attack are not capable to be predicted by doctors even due to the absence of specialists and more expansion of wrong diagnosed cases that bring required fabrication of an effective cardiovascular diseases prediction framework. This need prompted innovative work for new medical data mining strategies and different machine learning methods. The fundamental goal for this work is to recognize those key patterns and more features starting with that medical data utilizing classification algorithms to select the practical important attributes for silent heart attack diagnosing. The

utilization for recurrent neural system will further improve that exactness of the outcome and at the same time the usage of such a system is not unprecedented, those existing frameworks because of drawbacks and more undiscovered possibilities from silent heart strike. This paper means on delivering this process execution about inventive features to create more compensated systems.

Section 2 gives Literature Review, where the Proposed System and architecture is explained in Section 3, the Section 4 & 5 gives experimental results and conclusions about the work.

2. LITERATURE REVIEW

A novel CNN-MDRP [1] is been indicated through which high contaminated disease is most likely anticipated. A recurrent structure is developed to capture contextual data. The novel deep learning development is displayed in Bi-CNN-MI paraphrase identification (PI) [2]. Those PI thinks over two penalties on diverse stages of granularity. They choose Assuming about the reward which has same importance. Those factors of each model need to be upgraded to PI. Use of vernacular displaying assignment is to be addressed through the absence in preparing the data. They bring related investigation for algorithms such as decision Tree, bayes algorithm, SVM & nearest neighbor [3]. These calculations need to be used by the end in foreseeing bundle enrolment for majority of the data illustrations. They provide for a relative examination of distinctive calculations. In data mining they remove those disguised prescient data including the broad database. The ability of EHR [4] may be for setting up the new patients toward revealing that dark ailment connection. On EHR and the data mining techniques are about to grow by moral, real and specific reasons might prevent those deliberated testimony. Those tele-health administrations would be utilized to be known as the tele-health warning administrations. They need to be a large portion a component used similarly as a part of metropolitan urban communities. Because of tele-health administrations those patients might get aid undoubtedly [5]. A quick incremental in the tele-health system has gotten different ,methods like cloud computing and big data that has dynamic programming to create idle arrangements with the goal of information sharing system that takes care about transmission probabilities of planning imperatives and augmenting system limits.

To A content conclusion examination for jointed CNN & RNN design [6], brings those upsides about both in span

grained neighborhood for characterizing highlights which would be made toward CNN and long-separate dependencies discovered by method for the RNN. Those territorial constant infection needs been concentrated. Consideration need been paid ahead both organized also unstructured information. It uses a max-pooling layer that naturally judges, which expressions assume a paramount part in content order on catch enter parts previously, writings [7]. That majority of the data holds over qualities with missing majority of the data esteems need aid discriminating over upgrading choice - making system of cooperation. Those Taking in transform with respect to every event is indispensable similarly as it might hold numerous a couple exceptional Taking in. There need aid separate systems with manage lost majority of the data previously, choice tree Taking in. Those suggested credit count will be In view of the inherited computation that uses space esteems for that property as pool of courses of action. Survival of the fittest may be those introduce of hereditary calculation. The wellness fill in may be grouping precision from claiming a event with credited an impetus on the decision tree. The overall chase framework used byan and only hereditary calculation may be depended upon will get ideal result [8]. Our method viably combined range taking in around the medical characterization about both maladies Furthermore EHRs under a data determined methodology. Exploratory results with respect to a certified dataset from a recuperating office exhibited the viability of our suggested methodology. Their approach united undertaking relatedness, i. e., how each sickness relates for others, done a suitableness way, which provoked An progress in the prescient execution. Those joining about zone taking in over the restorative request for EHRs might have been also capable. Besides, the following impacts of the investigations of the illness particular prescient highlights not Exactly held discoveries enduring with existing medicinal territory learning, yet Moreover conveyed a couple hypothetical proposals [9]. Their system and effects Might a chance to be capable will overhaul that perception from claiming illness specific settings and moreover will upgrade that prescient execution done mortal sin displaying done intense doctor's facility mind. On [10,11] it displays a straightforward, yet skilled adjustment of the basic RNN (SRN) architecture, the accuracy RNN (CW-RNN), over which those concealed layer will be separated under divided modules, each taking care of commitments at its identity or specific worldly granularity, making calculations barely toward its proposed clock rate. Concerning illustration contradicted to settling on those standard RNN models more psyches boggling, CW-RNN lessens the amount of SRN parameters, enhances those execution in a far-reaching way in the undertakings tried, and accelerates the organize assessment. The framework may be illustrated for preparatory trials including three errands: sound sign generation, TIMIT talked expression classification, the place it beats both SRN and LSTM systems, also web handwriting recognition, the place it beats SRNs.

3. PROPOSED SYSTEM

Presented system utilized for forecasting of cardiovascular failure to meet the ideal precision in the outcomes. As found in the writing study, the AI methods utilized are pushing the precision till a specific breaking point. Besides, the issue with the current coronary episode forecast framework is the employments of traits. The credits to be chosen for the expectation of cardiovascular failure are the ordinary ones and consequently the outcomes are creating incorrectly results many-a-times. The proposed model expects to separate the best possible traits from the datasets which will upgrade the accuracy of the forecast. It will likewise give the clients legitimate analysis so the client comprehends the issue well absent a lot of trouble.

The undertaking separates itself by bridling the forces of both Deep learning and information mining. The paper proposes a framework, with a solid expectation calculation, which actualizes ground-breaking order ventures with an extensive report age module. The task expects to execute a self learning convention to such an extent that the previous contributions of the illness results decide the future prospects of the coronary illness to a specific client. The proposed model utilizes solid preprocessing instruments so the grouping and forecast don't show any blunders identifying with the dataset. A colossal no. of preparing sets will be utilized to make the expectation increasingly precise. Not exclusively does the datasets yet in addition the ascribes to be utilized are chosen mulling over the different significant boundaries and properties.

Actualizing the in advance of referenced model meets the objective of building up a framework with expanded exactness of evaluating the patient getting cardiovascular failure rate while tending to the disadvantages of the current framework The model which is proposed for Heart Attack Prediction System is made utilizing Deep learning calculations and approach. The proposed Flowchart for work is appeared in Fig.1

3.1 Data Collection

Using this method, we can extract the data of any kind of disease for diagnosing.

3.2 Pre Processing

In estimating the disease of heart, it pre-requisites to normalize & pre-process voluminous amount of information from the dataset.

3.3 DataMining

The predictive and effective system could be attained by strong pre-processing application or diminishment schemes of feature data on database of heart-disease like PCA, analysis of linear discriminant, and projecting randomly. Moreover, these schemes eradicate unwanted and unrequired data from huge former classification data. Hence, it has been introduced some of the schemes prominently into the models of data mining for lessening the space of memory, data utilization in database or server.

3.4 Evaluation & KDD

Moreover, it would allow further heart disease estimation process in optimal accuracy with better time.

Figure 1: Data Mining Procedures

4. SOURCE OF DATA

The term dataset could be signified as gathering of objects data that are possessing divergent amount of features of attributes, which handle required item attributes. Here, we attained 75 attributes of medical along with 303 records from UCI repository as stated in contribution [3]. Further, after 320 records utilization for 12 attributes of medical, 7 of them possess divergent qualities, where as 6 of they have stable values.

This system examined heart disease estimation with huge amount of attributes that are input. The terms related to medical are gender, age, cholesterol features such as 12 of attributes have been estimated the scope of heart disease patients. Moreover, the dataset comprises 12 of attributes, where its explanation has been stated in the following:

Table 1: Dataset contains 12 attributes and its description

Name	Type	Description
AGE	NORMAL	NO OF YEARS
GENDAR	DISTINCT	FEMALE: 0 MALE :1
Testbps	NORMAL	BLOOD PRESSURE BP in mmHg
ACCURATE HEART BEAT	DISTINCT	YES: 1 NO: 0
Thal	DISTINCT	NORMAL: 3 DEFICIENCY(FIXED): 6 DEFICIENCY (REVERSE): 7
Cp	DISTINCT	TYPES: Heart Attack: 1 Angina: 2 Aortic: 3 Pericarditis: 4
ST/HR Slope	DISTINCT	ST/HR Slope UP: 1 FLAT:2 DOWN:3
Chol	NORMAL	Cholesterol levels per deciliter (mg/dL)
Fbs	DISTINCT	Normal blood sugar>120 mg/dl: 1=NORMAL 0=DIABETIC
Old peak ST depression	NORMAL	0-6.2
Ca	NORMAL	Number of major vessels colored by fluoroscopy that ranged between 0 and 3.
Class	DISTINCT	Diagnostic classes:0= No presence 1=least likely to have heart disease 2=>1 3=>2 4=More likely have heart disease

4.1 ARTIFICIAL NEURAL NETWORKS

ANNs is the form of Neural Networks. A biological oriented network will be formed to predict the diabetics. The depiction of the neural network is shown in below figure 2:
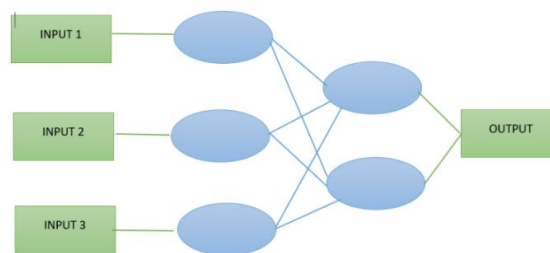

Figure 2: Representation of Neural Network

ANN must to be modified for each application else it leads to poor presentation. The black box method will be demerit discovered in ANN. ANNs is used for breast malignancy prediction in [14]. In [15] the authors applied ANN on 2 diverse breast malignancy datasets. Both of these datasets uses morph metric attributes. An improved ANN model [16] is used. Back propagation is used to prepare the systems [17, 18, 19].

4.2 Hybrid RNN

This deep learning approach need basic framework with an inherent sentiment circle permitting it should go about as a determining motor. RNN generally may be a standard neural system with an extra concealed state the place the concealed state impacts that neural system yield. Those stowed away state may be updated once every enter venture. It will be a model which cannot just figure out neighborhood and transient dependencies to information as well as might suit variable succession lengths. On RNN the yield of a layer will be included on the next enter layer Furthermore nourished again under those same layers. Dissimilar to feed-forward networks RNN might get grouping about qualities Likewise inputs and it can generate grouping of qualities as outputs. In any case for RNN issue for vanishing gradient may be more awful in this way it prompts exponentially little gradients Also rot for data through the long time

That pre-processed information will be provided for similarly as information at those enter layer. The expressions embedding layer changes over the data information under thick vector representational. For that assistance from claiming Taking in GRU layer those vector values need aid transformed. The SoftMax actuation work display in fully associated layer serves for classifying under coronary illness alternately non-heart diseases.
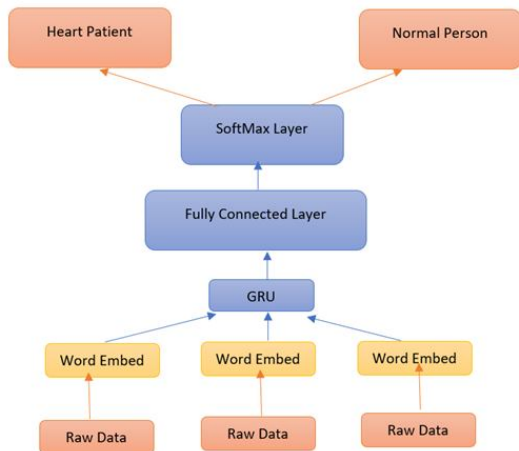
Figure 3: Data collecting from Heart patient and Normal Person

5. RESULTS

RNN would be executed in python & the streamlining utilizing the ANO, the library would perform functions once a GPU. Information measurements were 12 for each quality & 150 hidden data layers have been decided aimed at the RNN-HYBRID. 250 records have been utilized aimed at preparation situated Furthermore 120 records were utilized to trying situated. That model might have been prepared for 100 epochs & clump size similarly as 10 Also ADA delta might have been utilized similarly as those streamlining algorithm. At each epoch, "Validation AUC" would be ascertained utilizing the acceptance set, & assuming that it may be those best "Validation AUC" In this way far, those test set would be used to figure "Test AUC". The model with those best "Test AUC" might have been spared during those conclusions of the preparing which went should a chance to be 92.12%.

Accuracy of the model = 92.12%

Table 2: Accuracy of the various Techniques

Utilized strategy	Accuracy
Decision Tree &CNN [5]	81.52%
SVM [6]	85.02%
NB [7]	87.23%
KSOM [8]	87.99%
C4.5 MAFIA- [9]	89.23%
DBN [10]	89.90%
HYBRID RNN-	92.12%

The proposed algorithm has been implemented and the energy consumed is given in below Table 2 and graph. As shown the figure 4 is comparing various algorithms with energy levels.

Table 3: Energy Consumption

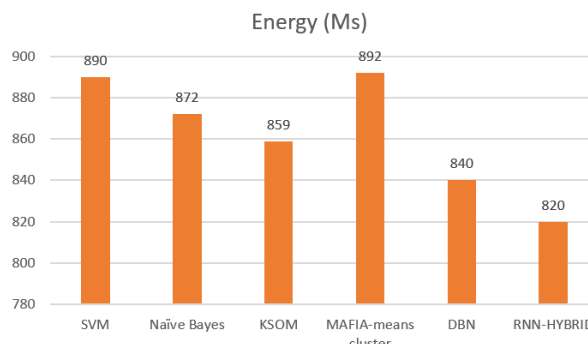| Technique used	Energy (Ms)
SVM	890
Naïve Bayes	872
KSOM	859
MAFIA-means cluster	892
DBN	820
RNN-HYBRID	820

Figure 4 : Comparative Analysis

6 CONCLUSION

Data Mining plays a vital role in the industry of health care, most effectively in predicting different and enormous type of heart diseases. The diagnosis is most widely being utilized in predicting diseases which helps in medical diagnosing. In abstract, no single method of data mining model is available to solve problems in health care. The proposed hybrid model obtains highest accuracy between classifiers which is much significant in medical diagnosing. In future work is done in this field so to improvise the treatments & the patient's lifetime by properly maintaining and analyzing the health sector data for directions is to improve the predictions utilizing hybrid models.

REFERENCES

1. Min Chen, Yixue Hao, Kai Hwang, Lu Wang, and Lin Wang, "Disease Prediction by Machine Learning over Big Data from Healthcare Communities." IEEE transaction, 2017, pp 8869-8879.
2. W. Yin and H. Schutze, "Convolutional neural network for paraphrase identification", in HLT-NAACL, 2015, pp. 901-911.
3. Seema sharma, Jitendra Agarwal, Shikha Agarwal, Sanjeev Sharma, "Machine Learning Techniques for Data Mining :A Survey", in Computational Intelligence and Computing Research, IEEE International Conference on. IEEE, 2013, pp.1-6.
4. Jensen PB, Jensen LJ, Brunak S, "Mining electronic health records: towards better research applications and clinical care, " Nat Rev Genet.2013 Jan; 14(1):75.
5. L. Qiu, K. Gai, and M. Qiu, "Optimal big data sharing approach for tele-health in cloud computing",

in Smart Cloud (Smart Cloud), IEEE International Conference on. IEEE, 2016, pp. 184-189.

6. Siwei Lai, Xu Kang Liu, Jun Zhao, "Recurrent Convolutional Neural Networks for Text Classification", in proceeding of the twenty-ninth AAAI Conference on Artificial Intelligence 2015.

7. Xingyou Wang, Weijie Jiang, Zhiyong Luo, "Combination of Convolutional and Recurrent Neural Network for Sentimental Analysis of Short Texts", International Conference on Computational Linguistics: technical papers, 2016, pp 2428-2437

8. Dipak V. Patil, R.S. Bichkar, "Multiple Imputation of Missing Data with Genetic Algorithm based Techniques", IJCA Special issue on Evolutionary Computation for Optimization Technique, 2010.

9. Ying Wen, Weinan Zhang, Rui Luo, Jun Wang, "Learning text representation using recurrent convolutional neural network with highway letters", NeuIR 16 SIGIR Workshop on Neural Information Retrieval, July 21, 2016, Pisa, Italy.

10. N. Nori, H. Kashima, K. Yamashita, H. Ikai, and Y. Imanaka, "Simultaneousmodeling of multiple diseases for mortality prediction in acute hospital care", in Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2015.

11. J. C. Ho, C. H. Lee, and J. Ghosh, "Septic shock prediction for patients with missing data", ACM Transactions on Management Information Systems (TMIS), vol. 5, no. 1, p. 1, 2014.

12. Kanchan M. Tarwani, Swathi Edem, "Survey on Recurrent Neural Network in Natural Language Processing", International Journal of Engineering Trendsand Technology (IJETT), Vol. 48 no. 6, June 2017.

13. JanKoutnik, Klaus Greff, Faustino Gomez, Jurgen Schmidhuber, "A Clockwork RNN", IDSIA, USI SUPSI, Manno Lugano, CH-6928, Switzerland.

14. R.Vijaya Kumar Reddy, K. Prudvi Raju, M. Jogendra Kumar, CH. Sujatha, P. Ravi Prakash, " Prediction of Heart Disease Using Decision Tree Approach", International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 6, Issue 3, March 2016.

15. NetworksDimitrios H. Mantzaris, George C. Anastassopoulos and Dimitrios K. Lymberopoulos, " Medical Disease Prediction Using Artificial Neural", BioInformatics and BioEngineering, 2008. BIBE 2008. 8th IEEE International Conference on 08 December 2008.

16. Youn-Jung Son, Hong-Gee Kim, Eung-Hee Kim, Sangsup Choi, and Soo-Kyoung Lee, "Application of Support Vector Machine for Prediction of Medication Adherence in Heart Failure Patients", Health Information Research, Vol.16, 2010.

17. K. Sai Prasad, Dr. S Pasupathy, "Deep Learning Concepts and Libraries Used in Image Analysis and Classification", TEST Engineering & Management, Volume 82, ISSN: 0193 - 4120 pp. 7907 – 7913

18. N Chandra Sekhar Reddy Purna Chandra Rao Vemuri; A Govardhan, "An Implementation of Novel Feature Subset Selection Algorithm for IDS in Mobile Networks", International Journal of Advanced Trends in Computer Science and Engineering (IJATCSE) ,Vol.8, No. 5, ISSN 2278-3091, pp 2132-2141.

19. N.Thulasi Chitra, K.Pushpa Rani, Roja, "A Credit Card Fake Detection System using Image Cryptography" International Journal of Recent Technology and Engineering (IJRTE) with ISSN: 2277-3878 (Online), Vol. 7 Issue-6, March 2019