International Journal of Advanced Science and Technology
Vol. 29, No. 8, (2020), pp. 1772-1782

Predictive & Preventive Healthcare: A Granular Distributed Model Utilizing An Empirically Effective Activation Function

B.Venkateswarla Chowdary [1], Dr.Y.Radhika [2]
[1]*Research Scholar,*[2]*Professor*
Dept of CSE, GITAM (Deemed to be University), Visakhapatnam, A.P India
[1]*bvchowdary@gmail.com,* [2]*yradhikacse@gmail.com*

Abstract

Healthcare analytics is most confidently considered as an specific domain that has interconnected healthcare analytics for numerous areas, as well as precautionary health, well-being and unwellness management. In particular to chronic diseases management, affected populations in several disease classes, the field of healthcare analytics assists target the customized management procedures and conventions that may alleviate the chronic disease, in addition as hamper the onset of affiliated medical conditions. Due to increase in size of patient related data in hospitals over a decade, development of fast information processing systems for this massive, noisy and fuzzy natured data is the need of the hour. The prototype design and analysis proposed in this paper aims to explore various relations among patient attributes, disease risk factors which can be mapped to built automated disease categorization and preventive healthcare systems. This paper presents the results obtained after performing experimental simulations on a standard healthcare related dataset. To present the advantage of the proposed model, the evaluated results are being compared with some existing approaches implemented in the past years.

Keywords: *Chronic disease, Preventive health care, neural network, Machine learning, supervised learning, Medical applications.*

1. Introduction

Innovations in the computing mechanics have reformed health care in recent past years. The analytical kind of the reasoning has not solely modified the manner within which data is possessed and reserved but has additionally assert associate degree measure progressively that plays necessary role within the administration and conveyance of healthcare prototypes. In particular, knowledge investigation has risen as a promising tool for resolution issues in numerous healthcare-related regulations. From a man of science and practicing perspective, a significant challenge in health care is its interdisciplinary nature. The sector of health care has typically observed advances in various disciplines eg. Databases, data processing, data resurgence, image process, medical domain researchers, and healthcare professionals [1][3]. Whereas this knowledge base not only adds to the affluence of the sector, it also reckons to the confronts in creating important advances. Computer analysts are sometimes not competent in domain- peculiar [4-7] medical ideas, whereas medical professionals and researchers even have restricted liability to the information analytics space. Recently, many researchers have carried out studies that reflected that by assimilating this supportive technologies, these are able to scale back fatality rates, aid prices, and medical aggravations at varied hospitals.

1.1 Healthcare Predictive Analytics: Clinical Analysis Methods

Clinical prediction forms an essential part of modern aid. Many prediction models are extensively investigated and further with success, setup in clinical aspects. Such kind of models have created an amazing impact in particular to designation and cure of diseases. Most of the supervised learning strategies that are used for clinical oriented prediction scenarios fall in three specific categories: (i) applied statistical strategies like, linear regression, along with Bayesian prototypes; (ii) strategies in machine learning as well as data processing like decision trees conceptualization and neural networks; (iii) Endurance models that intend to anticipate endurance results. Those procedures focus on finding the basic connection between covariate factors, which are called qualities, and a needy result variable. The decision of the model to be utilized for a particular guide drawback basically relies upon the results to be anticipated. There exist varied forms of the prediction blueprints which are supposed to be planned within the literature for supervising such type of the outcomes, a number of the foremost common outcomes include binary as well as continuous forms, alternative less normal structures are absolute and ordinal results. In addition to this, there exist totally different prototypes planned to handle the survival scenario outcomes wherever the motive is to predict the specific time of incidence of a specific event of particular interest. These survival prototypes are majorly studied within the aspects of clinical knowledge analysis in the perspective of predicting the average survival time of the patient.

1.2 Contribution Highlights

A prototype design is presented in this paper which aims to explore various relations among patient attributes, disease risk factors which can be mapped to build automated disease categorization and intelligent healthcare systems. The experimental evaluation is carried out on the standard Breast Cancer Prognostic dataset. The obtained experimental results are compared with few standard existing models.

1.3 Organization order of the Paper

Rest of the paper is structured as follows - Section 2 discusses about some background preliminaries. Section 3 is devoted to related work summary in this domain. The proposed model procedure is presented in section 4. Experimental analysis and obtained simulation results along with comparative summary are mentioned in Section 5. Finally, Section 6 of the paper present conclusions.

2. Background Preliminaries

This section presents some significant definitions e.g. machine learning, different types of learning, feature engineering, dimensionality and heterogeneity of data, bias and variance, discrete and continuous data, confusion matrix etc.

2.1 Machine Learning

Machine Learning (ML) is the point at which a system has been instructed to perceive designs by furnishing it with information (or preparing information) and a certain computation which helps in understanding that information. We allude this procedure of learning as preparing and the yield of this procedure is alluded as a model. The model is taken care of with new information (or test information) and it can reason about this new data dependent on what it has recently realized. AI models decide a lot of rules utilizing tremendous measures of registering power that a human cerebrum would be unequipped for preparing. The more information and AI model is taken care of, the more perplexing the standards and the more exact the forecasts. To sum up, the objective of AI is to comprehend the structure of information with the goal that precise forecasts can be made dependent on the properties of that information. While a factual model is probably going to have an inborn rationale that can be comprehended by a great many people, the standards made by AI are regularly outside human ability to understand on the grounds that our minds are unequipped for processing and dissecting colossal informational collections.

2.1.1 Supervised learning

Deals with making the function learn from the available training part of dataset. A supervised learning procedure exploits the available training data part and makes an inferred natured function, that can be then exploited further for mapping the new ones. Multiple supervised learning algorithms are available such as Support Vector Machines, Neural Networks and Naive Bayes categorizers etc.

ISSN: 2005-4238 IJAST

2.1.2 Unsupervised learning

Deals with un-labelled data without taking any previously defined dataset for model training. Unsupervised learning can be thought of as a potent tool for identifying patterns and trends and analyzing available data. There are various approaches used by unsupervised learning e.g. K-means clustering, hierarchical accumulation, self- formulating maps etc. Numerous significant learning types are - Active learning, Kernel-based learning, Transfer learning, Distributed learning, Association rule learning, Inductive logic programming, Reinforcement learning, Similarity and metric learning.

2.1.3 Feature Engineering

Feature Engineering is a procedure of changing crude information into an attribute vector which helps in expanding the prescient intensity of computations. It is the most significant craftsmanship in AI which makes the immense contrast between an optimal model and a terrible model. For instance - Suppose, the dataset incorporates the scope, longitude, and other information with the given name "Cost of House". The assignment is to foresee the cost of the house here. The scope and longitude are not of any utilization in the event that they are separated from everyone else. Thus, here the crossed section feature building is utilized. Steps to do feature engineering are:-

• Conceptualize features.

• Create attributes.

• Check how the attributes function with model.

• Start again iteratively from first till the attributes work utterly.

2.1.4 Dimensionality and Heterogeneity of data

Dimensionality in machine intelligence alludes to what number of attribute variables are available in the dataset. At the point when the dimensionality expands, the volume of the space increments so quickly that the accessible information becomes meager. The scourge of dimensionality reveals to us that evaluating a few amounts gets more diligently as the number of measurements of informational index increments - as the information gets massive or more extensive. For instance, social insurance information is famous for having immense measures of factors (for example circulatory strain, weight, cholesterol level). Heterogeneity is a word that means decent variety. A study hall comprising of individuals from heaps of various foundations would be considered having the nature of heterogeneity. The prefix hetero - signifies "other or unique". A heterogeneous populace or test is one where each part has an alternate an incentive for the trademark you're keen on. For instance, patients are commonly an exceptionally heterogeneous populace as they vary with numerous variables including socioeconomics, analytic test outcomes, clinical chronicles, and so on.

2.1.5 Bias and Variance

Bias implies how far away the expectations are from genuine qualities. The error because of predisposition is taken as the contrast between the normal (or normal) forecast of the model and the right worth which is to be anticipated. As there is just one model, discussing expected or normal forecast esteems may appear to be somewhat peculiar. In any case, the entire model structure procedure could be repeated on different occasions. Each time an alternate model could be accumulate new information and run another examination making another model. Because of haphazardness in the basic informational indexes, the subsequent models will have a scope of forecasts. Therefore, predisposition marks how far away when all is said in done these models' expectations are from the right worth.

ISSN: 2005-4238 IJAST

2.1.6 Discrete and Continuous Data

There are two kinds of quantitative information, which are additionally alluded to as numeric information: consistent and discrete. As a general rule, the count is a discrete entity, and estimations are ceaseless. Discrete information is a tally that can't be made more precise. Commonly it includes numbers. For example, the quantity of kids (or grown-ups, or pets) in your family is a piece of discrete information, as it is checking entire unbreakable elements: it is absurd to expect to have 2.5 children or 1.3 pets.

Constant information, then again, could be separated and decreased to better and better levels. For instance, the stature of children can be estimated all the more absolutely, for example, scale meters, centimeters, millimeters, and past so stature is ceaseless information.

2.1.7 Confusion Matrix

A confusion matrix is an N × N framework, where N is the count of classes being anticipated, used to gauge the exhibition of an arrangement model (or "classifier") on a lot of test information for which the genuine qualities are known. It has four distinct mixes of anticipated and real qualities as displayed in figure 1.

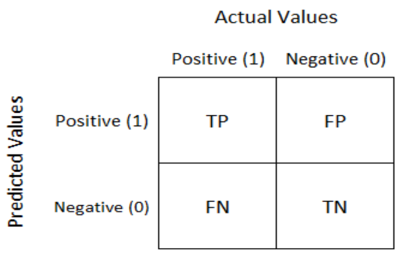

Fig.1 - Confusion Matrix

Confusion Matrix is extremely helpful in calculating Recall, Precision, Specificity and Accuracy.

3. Related Work

Parisa N. et al. [1], in their work, presented the application usecase of Multilayer Perceptron model specific to Neural Networks along with Support Vector Machines in the classification scenario of healthcare data. Leili T. [2] proposed a prototype for prediction of outlast and metastasis in the breast cancer patients. Sara K. et al. [3] performed the cluster reasoning for detecting the patterns of prescription use from consistently enthralled medical data. Mehrbakhsh N. et al. [4] had given a systematic strategy for disease prognosis and forecast utilizing AI methods. Kaur et al. [5] given a predictive modelling and analytics methodology for diabetes. Li et al. [6] had given physical action and danger of cardiovascular sickness Meta examination. Xin Li et al. [7] had given collaborative filtering-enhanced deep learning approach for heart disease prediction.

Zulfat M. et al. [8] start to end profound structure for sickness named entity identification utilizing internet-based social information. Abdulhamit S. et al. [9] proposed IoT based portable social insurance framework for human movement identification. Hongya Lu et al.[10] given a binary tree complex wavelet change based convolutional neural system for human thyroid clinical image segmentation. Sandeep P. et al. [11] Given a clinical IoT based system for eHealth care. SP Rajamhoana et al. [12] have performed examination of neural systems-based coronary illness prognosis framework. Amit W. et al. [13] given ECG characterization and prognostic methodology towards the customized social insurance. Haishuai W. et al. [14] Proposed a technique for foreseeing emergency clinic readmission by means of cost-sensitive deep learning.

ISSN: 2005-4238 IJAST

4. Proposed Model

4.1 Detailed Algorithmic Procedure

1. BEGIN PROCEDURE

2. Let, $\alpha_X[]$ and $\alpha_Y[]$ are input matrix and output matrix vectors respectively.
   Combination of α_X []: α_Y [] act as decision system for healthcare clinical analysis.
   Initialization process: Assign the random values to neural weights and biases correspondingly inside the network.
   Define -
   $\beta_H \rightarrow$ weight mat representation of hidden(intermediate) layer
   $\delta_H \rightarrow$ bias mat representation of hidden(intermediate) layer
   $\beta_O \rightarrow$ weight mat representation of outcome layer
   $\delta_O \rightarrow$ bias mat representation of outcome layer

3. The linear mapping will be performed as -
   IN_H = Mat dot product (α_X [], β_H) + δ_H.

4. Utilize Leakier Activation function {
   def: as f(x) = max (0, x)
   threshold as small amount of negative slope (0.01 approx.)
   fn. computes, f(x) = 1(x < 0) (ax) + 1(x >= 0) (x)
   here, a is small constant}
   perform non-linear mapping as -
   $\text{Act}_{LeakyReLU\ (H)}$ = LeakyReLU (Hidden Layer Input)

5. Outcome layer Input ← Mat dot product ($Act_{(H)} * \beta_O) + \delta_O$
   Outcome ← LeakyReLU (Outcome layer Input)

6. In this step, prediction extent is correlated with the absolute output and gradient error calculation is carried out as -
   Err = α_Y [] - Absolute Output

7. Cal, $\text{descent}_{Outcomelayer}$ = derivatives LeakyReLU (Absolute Output)
   $descent_{(H)}$ = derivatives_LeakyReLU ($Act_{(H)}$)

8. $Delta_{Outcome}$ ← Err * $descent_{Outcomelayer}$

9. Err Backprop: $Err_{(H)}$ ← Mat dot product ($\text{Delta}_{\text{Outcome}}$, β_O^T)

10. Measure: $Delta_{(H)} = Err_{(H)} * descent_{(H)}$

11. Neural weights updation -

    $\beta_O \leftarrow \beta_O$ + Mat dot product ($Act^T_{(H)}$, $Delta_{Outcome}$) * learning rate
    $\beta_H \leftarrow \beta_H$ + Mat dot product ((α_X [])^T, $Delta_{(H)}$) * learning rate

12. Biases updation –
    $\delta_H \leftarrow \delta_H$ + sum ($Delta_{(H)}$, axis=0) * learning rate
    $\delta_O \leftarrow \delta_O$ + sum ($Delta_{Outcome}$, axis=0) * learning rate
13. END PROCEDURE

ISSN: 2005-4238 IJAST

The clinical healthcare decision support systems are designed to provide assistance and analytical abilities to the clinicians decision making through patients records such as diagnosis and treatment parameters attributes. Steps (2-5) are forward information propagation and steps (6-12) act as backward information propagation.

4.2 Analysis

This section summarizes some critical analysis of the proposed model procedure.

Observation (i): In the proposed procedure, non-linear functional plot is adopted, which is advantageous i.e. with this, model can learn and map almost any inconsistent complex things as input to output.

Observation (ii): Back propagation based optimization strategy is employed as steps (6-12) in the pseudocode which is able to compute loss gradients.

Observation (iii): The activation function (Leaky ReLU) utilized in the proposed procedure gives benefits which are summarized as follows -

• Fast convergence rate

• Avoids"Vanishing Gradient Problem (VGP)", so it results better accuracy and improved overall model performance

• Gradients are less fragile during training and there is a scope to keep weight updates active

• Comparatively easier optimization

5. Experiments Analysis and Results Discussion

The experimental simulation setup, input dataset detail, obtained results and comparative analysis are discussed in this section.

5.1 Setup (Hardware and Software) details

Experimental setup consists of system environment as Ubuntu 16.04 LTS with 64 bit OS, 8 GB RAM, processor as Intel Core i7-860 @ 2.80 GHz clock speed, number of cores as 4 and total 8 threads available. H2O framework (v1.5.2) with Python version 3.7 is used for simulation and statistical computing purposes.

5.2 Dataset Overview

Dataset name: Breast Cancer Prognostic Data

Source: UCI ML Repository Attribute Characteristics: Real Number of Attributes: 34

Data Set Characteristics: Multivariate

ISSN: 2005-4238 IJAST

Number of Instances: 198

5.3 Obtained Results

The model performance summary in terms of obtained statistical parameter is given in table 1. The graphical representation is shown as Fig 2.

Table 1: Statistical Performance Results

Performance	Value
Accuracy	0.9876
AUC	0.99486
F1 Score	0.9722
GINI	0.9897
LOGLOSS	0.0917
Simulation Time	3172 seconds

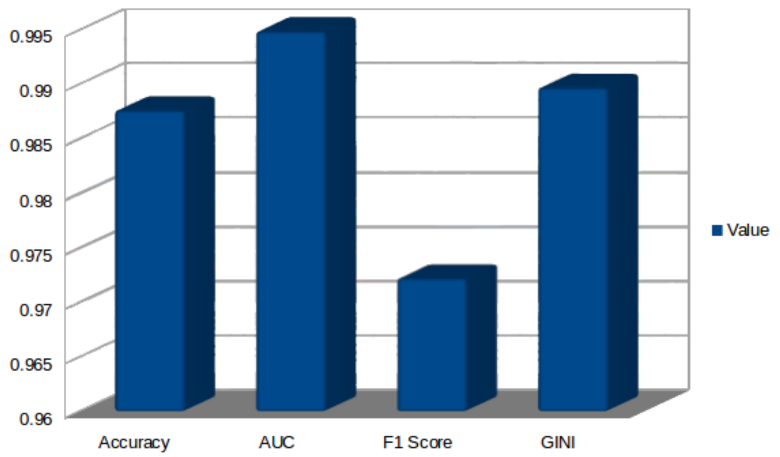

Fig.2 - Performance Parameter vs Value Representation

5.4 Comparisons with Existing Approach (es)

In table 2, the comparison of the obtained results is carried out with some significant existing methods - E. Zafiropoulos et al. [15], S.S. Shajahaan et al. [16] and J. Ivancakova et al. [17]. The comparative analysis shows that the simulation results outperform the existing methods. Graphical representation is shown in Fig 3.

Table 2: Comparison with Existing Approach(es)

Reference	Method	Prediction accuracy
E. Zafiropoulos et al. [15]	SVM with Gaussian RBF	89.28%
S.S. Shajahaan et al. [16]	CART	92.42%
J. Ivancakova et al. [17]	C4.5	95.61%
Proposed Method	Proposed Method	98.76%

ISSN: 2005-4238 IJAST

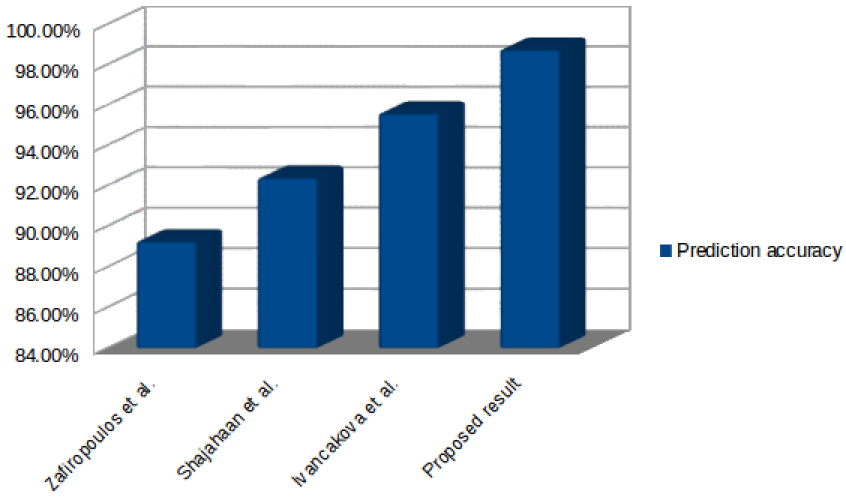

Fig.3 - Comparisons with Existing

6. Conclusion

The domain of information analytics has seen vital strides in recent past years that are attributable to hardware and package technologies, that have magnified the convenience of the information assortment method. The advancement of the healthcare sector has variety of challenges attributable to its knowledge base nature, privacy curtailments in information assortment furthermore, dispersal systems and the inalienably unstructured nature of the information. In specific cases, the information might have terribly massive volume, which surely needs real-time analysis and its insights. In certain cases, the information could also be arduous, which needs specialized retrieval as well as analytical methods. The advances in information processing technologies, which have facilitated the sphere of analytics, additionally create new confronts attributable to their potency in analyzing huge amounts of knowledge. This paper presents a computationally efficient approach for automated disease prediction along with the experimental evaluations and comparative analysis with existing methods.

References

[1] Parisa Naraei, Abdolreza Abhari, Alireza Sadeghian, "Application of Multilayer Perceptron Neural Networks and Support Vector Machines in Classification of Healthcare Data", FTC 2016 - Future Technologies Conference, 6-7 December 2016, San Francisco, United States, IEEE.

[2] Leili Tapak, Nasrin shirmohammadi-Khorram, Payam Amini, Behnaz Alafchi, Omid Hamidi, Jalal Poorolajal, "Prediction of survival and metastasis in breast cancer patients using Machine learning classifiers", Clinical Epidemiology and Global Health, S2213-3984(18)30182-9, 2017.

[3] Sara Khalid, M. Sanni Ali, Daniel Prieto-Alhambra, "Cluster Analysis toDetect Patterns of Drug Use from Routinely Collected Medical Data", IEEE 31st International Conference on Computer based Medical Systems, University of Oxford, UK, 2018.

[4] Mehrbakhsh Nilashi, Othman bin Ibrahim, Hossein Ahmadi, Leila Shahmoradi, "An Analytical Method for Diseases Prediction Using Machine Learning Techniques", Computers and Chemical Engineering, Volume 106, 2 November 2017, Pages 212-223.

[5] Kaur, H., Kumari, V., "Predictive Modelling and Analytics for Diabetes using a Machine Learning Approach", Applied Computing and Informatics (2018), doi: https://doi.org/10.1016/j.aci.2018.12.004.

[6] S. Velliangiri, P. Karthikeyan & V. Vinoth Kumar (2020) Detection of distributed denial of service attack in cloud computing using the optimization-based deep networks, Journal of Experimental & Theoretical Artificial Intelligence, DOI: 10.1080/0952813X.2020.1744196

[7] Xin Li, Juan Li, "Health Risk Prediction Using Big Medical Data - a Collaborative Filtering-Enhanced Deep Learning Approach", 2018 IEEE 20th International Conference on e-Health Networking, Applications and Services, 2018, pp.1-7.

[8] Praveen Sundar, P.V., Ranjith, D., Vinoth Kumar, V. et al. Low power area efficient adaptive FIR filter for hearing aids using distributed arithmetic architecture. Int J Speech Technol (2020). https://doi.org/10.1007/s10772-020-09686-y.

[9] Abdulhamit Subasi, Mariam Radhwan, Rabea Kurdi, Kholoud Khateeb, "IoT based mobile healthcare system for human activity recognition", 15th Learning and Technology Conference (L&T), 2018, pp. 29-34.

[11] Sandeep Pirbhulal, Wanqing Wu, Subhas Chandra Mukhopadhyay, Guanglin Li, "A Medical-IoT based Framework for eHealth Care", International Symposium in Sensing and Instrumentation in IoT Era (ISSI), 2018, pp. 1-4.

[12] SP Rajamhoana, C. Akalya Devi, K. Umamaheswari, R. Kiruba, K. Karunya, R. Deepika, "Analysis of Neural Networks Based Heart Disease Prediction System", 11th International Conference on Human System Interaction (HSI), 2018, pp. 233-239.

[13] Amit Walinjkar, John Woods, "ECG classification and prognostic approach towards personalized healthcare", International Conference On Social Media, Wearable And Web Analytics (Social Media), 2017, pp. 1-8.

[14] Haishuai Wang, Zhicheng Cui, Yixin Chen, Michael Avidan, Arbi Ben Abdallah, Alexander Kronzer, "Predicting Hospital Readmission via Cost-Sensitive Deep Learning", IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2018, Volume: 15, Issue 6, pp. 1968-1978.

[15] E. Zafiropoulos, I. Maglogiannis, and I. Anagnostopoulos. 2006. A support vector machine approach to breast cancer diagnosis and prognosis, Artificial

Intelligence Applications and Innovations (2006), 500-507.

[16] S. S. Shajahaan, S. Shanthi, V. ManoChitra. Application of Data Mining Techniques to Model Breast Cancer Data, International Journal of Emerging Technology and Advanced Engineering. vol. 3, issue 11, 2013, pp. 362-369.

[17] J uliana Ivancakova, Frantiek Babi, Peter Butka. Comparison of Different Machine Learning Methods on Wisconsin Dataset, IEEE 16th World Symposium on Applied Machine Intelligence and Informatics, Slovakia, February 7-10, 2018.