IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A Modern Approach to Predict Heart Disease using Neural Networks

[1]Murali Mohan T, [2]Yugandhar Bokka, [3]Karunakar K, [4]Archana G

[1]Professor, [2,3,4]Associate Professor

[1,2,3]Computer Science and Engineering

[1,2,3]Swarnandhra Institute of Engineering and Technology, Narsapur, WG(Dt), Andhra Pradesh,India.

[4]Swarnandhra College of Engineering and Technology, Narsapur, WG(Dt), Andhra Pradesh,India

Abstract: Obese individuals have a high possibility of developing Cardio Vascular Disease (CVD), which should be because of an alteration in the Autonomic Nervous System (ANS) .Heart Rate Variability (HRV) can be utilized to recognize the progressions in ANS movement. An HRV is a device to gauge the ANS action utilizing linear and non- linear HRV features. In this paper, we address the impact of weight on ANS utilizing HRV parameters. At first, two sets of twenty counts are taken; one set contains obesity and one more set is non-obesity regardless of gender, between the ages of 20 to 50. The measurably massive distinction between the two sets was noticed using the Independent t-test. Measurable consequences of the review show the sympathovagal imbalance because of reduced parasympathetic movement. To come up with exact outcomes, we executed deep learning techniques. This paper essentially represents two best profound learning calculations: RNN (Recurrent Neural Network) and LSTM (Long Short-Term Memory). The experimental results show that the LSTM expectation impact is superior to RNN in terms of recognizing CVD.

***Index Terms* - Deep learning, RNN, LSTM, CVD, ANS.**

I. INTRODUCTION

Heart disease currently affects a huge number of people worldwide and is a common place ailment in the human body. Every year, a large number of people in what are considered developed nations like the USA and India are also impacted by this disease. According to study, heart disease is caused bya prolonged period of living an unbalanced lifestyle and engaging in numerous types of bodily activity. .Based on symptom patients are worried and ready to take an immediate treatment at hospital with various tests and therapy's even though could not find right problem at that situation, still are paying more money to identify the correct disease. So it is lack of awareness to identify the heart disease [11–13]. In this research, we are trying to get good results with maximum accuracy to identify heart problem immediately.

In these period most of the people suffer with obesity. According to thedefinition of obesity, an excessive buildup of body fat causes chronic disordersincluding hypertension, cardiovascular disease (CVD), myocardial infarction(MI), and diabetes. Obesity and CVD are strongly correlated, according toresearch [1, 8]. According to the study's findings, obesity increases the risk ofCVD and autonomic activity imbalance [2]. The body's control system, theautonomic nervous system (ANS), is responsible for maintaining homeostasisin most cases. The internal organs, blood arteries, and glands are all regulatedby the ANS. It has two branches: the parasympathetic nervous system and thesympathetic nervous system (PNS). In contrast to the PNS, which controlsthe rest and digest response, the SNS mobilizes the body to supply energy forthe fight-or-flight reaction. The vagal nerve acts as a mediator between theautonomic nervous system and the heart, and HRV quantifies this effect [9].Heart rhythm alterations are impacted by even the slightest ANS changes.An electrocardiogram's HRV is a fluctuation in the RR interval (ECG). As aresult, HRV may be the most significant and non-invasive technique to researchhow obesity affects the ANS. Obesity substantially reduced HRV, which raisedthe risk of CVD [2, 3]. HRV characteristics, both linear and non-linear, can beused to model changes in the key organs that the ANS regulates in the body.Deep learning has currently gained widespread acceptance and has been successfully used in all facets of modern life. In the medical era, there arealso effective deep learning applications applied, such as illness risk predictionbased on medical data, as a diagnostic tool is accessible to help doctors detectthe condition [10]. In order to help clinicians diagnose patients, increase theprecision of pathological diagnosis, and enhance the quality of health care, deep learning may be applied to medical data. The two fundamental ideas of boththe RNN and LSTM deep learning algorithms are introduced in this article.

It is built a prediction model using two different deep learning networks. A unique RNN network is the LSTM network (Long Short-Term Memory). According to World Health Organization (WHO) recommendations, BMI is used to assess obesity. BMI is computed using the formula BMI = Weight (kg)/Height (m2) [4]. Subjects are classified as normal or non-obese if theirBMI is between 18 and 25 kg/m2, and as obese if it is greater than 30 kg/m2[5, 6]. The structure of this paper is as follows: Section II introduces the methodology, the obesity criteria, the statistical t-test, and the deep-learning algorithm; Section III discusses the results of the statistical and machine learning techniques; Section IV discusses the findings and conclusion; and Section VI discusses references.

II. PROPOSED METHOD

A heart disease sample with 13 columns and 7000 rows is acquired from Kaggle [7]. Each column has distinct qualities and influences heart disease either directly or indirectly. We tested and predicted using all features. This dataset has only one class label and remaining all attributes. All features based on age and gender, gender is considered as 1 and 2 for women and men respectively. Based on data set, we considered as smoking record, blood pressure, glucose, cholesterol, Physical activity, blood pressure. Some records are categorized witha 0 or a 1, where 0 and 1 denote absence and presence, respectively.

Table 2.1: Sample Data Set for Heart Disease

Id	age	gender	height	Weight	ap_hi	ap_lo	cholesterol	gluc	Smoke	alco	active	cardio
0	18393	2	168	62	110	80	1	1	0	0	1	0
1	20228	1	156	85	140	90	3	1	0	0	1	1
2	18857	1	165	64	130	70	3	1	0	0	0	1
3	17623	2	169	82	150	100	1	1	0	0	1	1
4	17474	1	156	56	100	60	1	1	0	0	0	0
8	21914	1	151	67	120	80	2	2	0	0	0	0
9	22113	1	157	93	130	80	3	1	0	0	1	0
12	22584	2	178	95	130	90	3	3	0	0	1	1
13	17668	1	158	71	110	70	1	1	0	0	1	0
14	19834	1	164	68	110	60	1	1	0	0	0	0

Recurrent Neural Network (RNN)

The major use of recurrent neural networks is the processing of sequential input. To handle long-standing behaviors more effectively, long- and short- term memory networks are utilized. Recurrent neural networks are made with hidden state usage in mind. Each concealed layer selectively forgets certain information while storing other information. In this regard, it is possible to extract the sequence's data properties and data changes. In addition to text processing, RNN is also often employed in the areas of sentiment analysis, speech recognition, machine translation, text generation, and video behavior detection. As a result, we used the RNN modeling approach in this article to forecast the risk of cardiac illnesses.

Long Short Term Memory (LSTM)

In the RNN layer, these connections are employed between neurons. Net- works have the ability to alter the before-and-after data sequence law, andit is simple to mine internal data sequence laws. Speech recognition and machine translation are two applications of sequence data processing where RNN is often utilized. However, there are certain issues with this model aswell. The problem of gradient disappearance or gradient explosion cannot be avoided when data is transferred backward, which restricts the processing of long-term relationships. By including several specialized computational nodes in the RNN's hidden layer, the LSTM network modifies the way gradientsare sent during back propagation, significantly slowing the issue of gradient disappearance or explosion.

Statistical analysis

The definition of a two-sample T-test is a statistical hypothesis testing method that compares the means of two distinct populations using two independent samples. When the standard deviations of the populations being compared areunknown and the sample size is limited, the two-sample T-test is employed. When two samples are independent and have normal distributions, the two- sample T-test is employed. You require two independent samples in order to use the two-sample T-test as explained in this work. The following formulamay be used to estimate the difference between the means of two sets:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

Difference in population means = Difference in sample means +/- T*standard error

$$(\bar{x}_1 - \bar{x}_2) \pm t^* \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

III. RESULTS AND DISCUSSION

A statistical t-test is used cholesterol and then comparing the averages of two separate populations (males and females).

Male population:
N1: 452
df1 = N - 1 = 452 - 1 = 451
M1: 1.37
SS1: 203.04
s21 = SS1/(N - 1) = 203.04/(452-1) = 0.45

Female population:
N2: 248
df2 = N - 1 = 248 - 1 = 247
M2: 1.39
SS2: 118.84
s22 = SS2/(N - 1) = 118.84/(248-1) = 0.48

The t-value is -0.36972. The p-value is 0.355851. The result is not significant at p 0.05.

Obesity is a heart condition that is brought on by a lifetime of poor eating, lifestyle choices, and environmental factors. Recurrent neural network and Long short - term memory heart disease risk prediction model as shown in Fig.1.

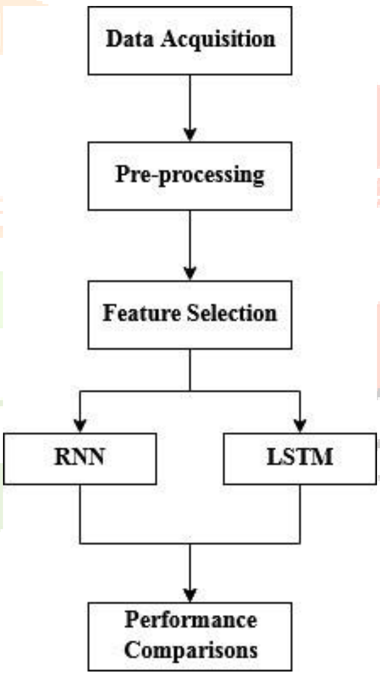

Fig. 1 Risk prediction model for Heat Disease based on RNN and LSTM

In this study, 7000 randomly chosen people served as experimental samples, of which 1400 served as test sets and 5600 as training sets. For training and prediction, LSTM models and RNN models are employed. Fig.2. displays the RNN and LTSM loss function's changing curve.

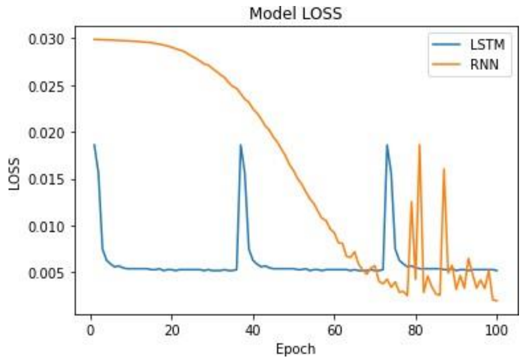

Fig.2: Risk prediction model for Heat Disease based on RNN and LSTM

In the training data set, the LSTM model converges more quickly than the RNN model, with smaller loss function values and greater accuracy. Fig.3. displays the findings as the prediction accuracy of the two models is compared.

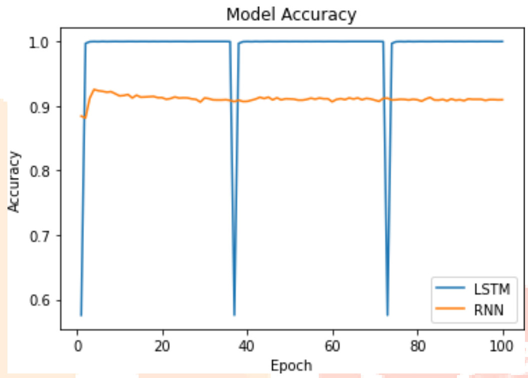

Fig.3: Risk prediction model for Heat Disease based on RNN and LSTM

IV. Conclusion

In this study, an RNN and LSTM-based risk prediction model for heart disease was developed. According to experimental findings, LSTM has a prediction impact that is 95% more accurate than typical RNN networks. In order to prevent, intervene with, and manage diseases, disease risk prediction refers to the identification of prospective hazards and trends of diseases. Finding possible risks and trends of illnesses before doctors detect them and then taking action to prevent and treat them are the ideal goals of disease risk prediction in medicine. Additionally, this model provides improved accuracy, and users may conduct tasks from a system and get awareness of their present heart disease condition by utilizing these models. In the future, we intend to employ Auto Encoders and Generative Adversarial Networks models to provide effective outcomes that will assist individuals reduce their chance of developing heart disease.

REFERENCES

[1] Hurt RT, Kulisek C, Buchanan LA, McClave SA. The obesity epidemic: challenges, health initiatives, and implications for gastroenterologists. Gastroenterol Hepatol (N Y). 2010;6(12):780792.

[2] Thayer JF, Yamamoto SS, Brosschot JF. The relationship of autonomic imbalance, heart rate variability and cardiovascular disease risk factors.Int J Cardiol. 2009; 141(2):122131.

[3] Rossi RC, Vanderlei LC, Gonalves AC, Vanderlei FM, Bernardo AF,Yamada KM, da Silva NT, de Abreu LC. Impact of obesity on autonomic modulation, heart rate and blood pressure in obese young people. Autonomic neuroscience. 2015;193:138-41.

[4] Billman GE, Huikuri HV, Sacha J, Trimmel K. An introduction to heart rate variability: methodological considerations and clinical applications. Front Physiol. 2015;6:55.

[5] World Health Organization: obesity and overweight. Fact sheet No.311,[updated June 2016]. Available from:http://www.who.int/mediacentre/factsheets/fs311/en/. AccessedDecember 13, 2016.

[6] Romero-Corral A, Somers VK, Sierra-Johnson J, et al. Accuracy of body mass index to diagnose obesity in the US adult population. Int JObes(Lond). 2008;3(6):959966.

[7] https://www.kaggle.com/datasets/sulianova/cardiovascular-disease-dataset?resource=download.

[8] Ama Moor, Vicky Jocelyne, et al. "Dyslipidemia in patients with acardiovascular risk and disease at the University Teaching Hospital of Yaounde, Cameroon." International journal of vascular medicine 2017 (2017).

[9] Bokka, Yugandhar, R. N. V. Jagan Mohan, and M. ChandhraNaik. Predictive Analysis for daASD Using Population based Incremental Learning. Journal of Engineering Science & Technology Review 14.3 (2021).

[10] Cui, Shiyue, et al. "Research on risk prediction of dyslipidemia in steel workers based on recurrent neural network and lstm neural network." IEEE Access 8 (2020): 34153-34161.

[11] S. Rathod, U. Chaskar and C. Patil, "Machine Learning Approach to Study the Impact of Obesity on Autonomic Nervous System using Heart Rate Variability Features," 2021 Seventh International conference on Bio Signals, Images, and Instrumentation (ICBSII), 2021, pp. 1-4, doi: 10.1109/ICBSII51839.2021.9445129.

[12] Maru, A. K. Sharma and M. Patel, "Hybrid Machine Learning Classification Technique for Improve Accuracy of Heart Disease," 2021 6th International Conference on Inventive Computation Technologies (ICICT), 2021, pp. 1107-1110, doi: 10.1109/ICICT50816.2021.9358616.

[13] R. Karim, W. Chengliang and M. Islam, "Comparative Study to Identify the Heart Disease Using Machine Learning Algorithms," 2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE), 2021, pp. 1097-1100, doi: 10.1109/ICBAIE52039.2021.9390032.