# A study on the methods of prediction of hypertension

Abhay Goyal[1], Durgesh Singh[2] and S. Sharanya[3]

*Abstract*— **Hypertension is an illness that often leads to severe and life threatening diseases such as heart failure, thickening of heart muscles, coronary artery disease and other severe condition if left unchecked. This survey provides an overview of the different research work that has been done in the treatment of hypertension. It has been said that hypertension prediction if employed in an efficient manner can save many lives. The paper aggregates the work done by many different researchers across multiple disciplines and domains to predict hypertension in humans, which has been a prime health issue in modern world. This paper presents how and what are the different methods that have already been employed and how can work further in this area.**

## I. INTRODUCTION

Hypertension is a health condition of a person that can lead to very severe illness live stroke, heart disease, and renal failure. The study in the hypertension is done through the machine learning and deep learning methods. According to the a major global organization, namely WHO in its recent survey has shown that high blood pressure causes one out of the eight deaths and hence Hypertension is considered the third leading killer in the world. There are more than a billion hypertensive patients around the word and around four million patients die every year. In the Middle Eastern region, cardiovascular disease and stroke are the main cause of death and illness. Hypertension has not only increased in these over the years but also has been a major cause of loss in revenue due to the loss of manpower.

Machine Learning is one of the most reasonable options that allows clinicians to identify better pattern in the data set that are sometimes possible are a harder task for humans and improve upon the working of medicine. Artificial Neural Networks (ANNs) are currently an upcoming field as is popular across a variety of domains. For example, ANNs is being used to predict the stock market, detect objects and perceive scenes. In the field of medicine in particular, ANNs are popular due to their ability to analyze medical images, design different kinds of medicines, and diagnose diseases. Medical systems made so achieve a high percentage of efficiency, accuracy, and reliability and minimize the harm patients can get into. This is why ANN is one of the most backed on and researched field. The applications and systems made integrate machine learning in order to provide methodical and accurate results that a doctor can use to diagnose and cure illnesses. Technologies formed provide timely interventions if made to work with data obtained early in a patients treatment cycle. They provide an insight into the techniques and different methods that have been researched on predicting the possibility of hypertension using the different features as mentioned further in the paper. Hypertension has become a priority for healthcare providers in the era of high burnout rate and workload.

This paper gives a survey of papers from various disciplines ranging from clinical psychology to Artificial Intelligence. The survey thus written has a great mix of all the different perspectives that have been studied during the course of the advancement which makes this paper unique. The paper provides exhaustive knowledge on how and what changes have taken place over the time.

The paper follows the following format, Section II is the Literature Survey. Section III is the Discussion on the survey done. Section IV describes the future work related to the papers surveyed, Section V is the Conclusion and Future Work

## II. LITERATURE SURVEY

A model intelligent hypertension prediction system was build with the help of machine learning and decision trees. By using the aforementioned technique the researchers were able to solve the well defined goal. The system to use decision trees to predict hypertension was proposed by [1]. The paper was proposed in 2016. The result illustrated did show the strength of the working model. It produced a result of 93.6%, which is a very good result by using the given data set. In this C4.5 algorithm was used create the decision tree and a total of 1100 data of people was used for the training and testing purpose. The solution presented in the paper to solve this was by developing a smart system of health related to the development of health risk. The prediction for hypertension treatment with machine learning approach and with additional prediction feature. The structure is divided into three parts. Learning process, in this section, the dataset provided and using the approach of decision tree algorithm. This process generates decision trees and produces a classification of the dataset that is inserted into the system, followed by the formation of decision rule. Decision rule is used to predict the diagnosis and prognosis of hypertension. The diagnosis of hypertension section predicts the diagnosis of hypertensive conditions based on the data entered. The information entered is a medical condition parameters obtained from the different measurement results taken during the diagonsis . The information generated in this process is then compared with a decision rule. A similar methods ids used for the proganosis. Similar to diagnosis, the

*This work was not supported by any organization

[1]Abhay Goyal is with Faculty Computer Science and Engineering, SRM Institute of Science and Technology, India

[2]Durgesh Singh is with Faculty Computer Science and Engineering, SRM Institute of Science and Technology, India

system uses decision trees to predict the different methods in which the decision can be made. The information generated in this process is done by comparing the data with a decision rule that is formed in the previous process. The whole system once finished was installed and it was found to have an accuracy of around 92%

In 2016, in a paper proposed [2], they used ANN (Artificial Neural Network) for the prediction of hypertension by the use of CPCSSN dataset. The model was used for a very large data set(185,371 patients and 193,656 controls). The result obtained was rather good considering the large data set used. The result obtained was 82% accuracy. There were 11 inputs with 8 hidden nodes and 2 outputs from this model. In this the data set was first preprocessed and then the data related to the 11 inputs are only entered in the training data or testing data. These include age of the patients, their gender, the BMI of the patient, systolic and diastolic blood pressure etc. The previous papers had found that a deep neural network with 20 nodes in the first hidden layer and 5 nodes in the second hidden layer result in the best accuracy of 92.85%. In [2], a 3 layer network with 11 input nodes, 7 hidden nodes and 2 output nodes which thus have a total of 91 weight values ((11*7) + (7*2)) was used. It was found that more hidden nodes, better mapping was the mapping of the input to the output and accuracy of the results increased. But it also had its own unpropitious effects of increased processing time and overfitting of data. In the dataset it was found that the patients may be suffering from one or more of the eight chronic diseases of interest to the researchers. These eight diseases included hypertension, diabetes, chronic obstructive pulmonary disease (COPD),etc. In the paper the patients set for control were individuals who were not suffering from any of the aforesaid conditions but did need medical treatments for various other reasons.

In 2018, the paper proposed by [3], They used waist to height ratio to predict the hypertension in people. The experiment was done in for a period of over 2.8 years in which 1718 participated and it was calculated that if the ration of the measured waist to measured height was greater than 0.5, then the chances of having hypertension increases by 4.5 times. It was exclusively a result for a small number of people and there might be many disadvantages for this. The baseline characteristics used chi-square test for examination. To approximately find which of the markers were the best predictor of hypertension the sensitivity, specificity and mean area was calculated under the AU-ROC of the baseline model made for BMI, WHR and WHtR, and their confidence intervals (CIs), using receiver operating characteristic curves in the group in which hypertension had newly developed. The ratio of waist to height was found to be a better method of calculating the hypertension probability rather than BMI. This is so because BMI does not differentiate between body fat and muscle mass, and hence such people have the same risk of such illnesses as others. The early detection of hypertension is very important for its control and prevention, in adults; but as, blood pressure is not tested very often in adolescents, it is difficult to monitor the their condition. This study has some limitations. First, the sample is low and hence the number of participants who developed hypertension after 2.8 years of follow-up were relatively small. Secondly, the population that was studied was limited to adults from rural areas, and, thus, does not represent the overall population. Third, there is selection bias present in the study done.

[4] uses PWV(Pulse Wave Velocity) and Blood pressure. The findings from the risk prediction model suggest that PWV predicts the progression of blood pressure and provide a valuable tool in hypertension risk prediction in young adults. The aim of the study was to predict the increase in blood pressure by reading the PWV and thus predicting hypertension. Blood pressure from the right brachial artery was measured in the sitting position after a 5-minute rest with a random-zero sphygmomanometer in 2007 and 2011. Linear regression was used to study the interdependence between PWV measured in 2007 and blood pressure measured which was absent in 2007 ,in 2011. Separate analysis were done on both of the populations (whole population and normotensive subpopulation). Logistic regression analysed relation of PWV measured in 2007 to the hypertension in 2011 which was absent in the study in 2007. Logistic regression was performed on approximately 1200 subjects who were normotensive in 2007. The regression model made also included traditional cardiovascular risk factors such as age, sex, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, triglycerides, glucose, insulin etc. All predictor variables besides age were made to a certain standard in order to make the variables effect sizes proportional to each other. Hence, the regression coefficients () in the linear regression models and the odds ratio in the logistic regression model indicated that there was effect of a 1-SD change in a predictor variable for any given dependent variable. There is no statistical significant interactions between sex, age, or BMI and hypertension in 2011. A major limitation of the present study was the lack of PWV data in 2011. Therefore, the researchers were not able to examine the effects of blood pressure and hypertension on PWV progression.

In a paper proposed by [5], the main hypertension point was Intracranial Hypertension. By the study of Arterial Blood Pressure(ABP) and Cerebral blood flow velocity was used to predict the Intracranial pressure(ICP). There were 36 patients suffering from traumatic brain injury. The Random forest classifier resulted in most stable and accurate prediction yielding a sensitivity of 69.1% and specificity of 78.3%. This paper was especially proposed for the patients that are suffering through brain injury to find intracranial Hypertension in the patients present. The prediction was done using beat detection and signal detection. Linear SVM, NB and RF classifiers were used. Bagging was used to add many different Random Forests and then implicit feature selection was done. 256 decision trees were used with the depth of 8 splits. The Random Forest classifier achieved the best performance in terms of sensitivity to the data and had a almost similar specificity to the other classifiers. Both the linear SVM and the NB classifier were outperformed by the RF in terms of sensitivity, which is the more relevant

statistical measure when screening patients. Hence the RF was a better method of classification

[6] is another paper which uses a host of machine learning algorithms and tries to compare between them. The dataset used in this paper, includes the information of approximately 7000 patients that had followed the physicians treadmill stress testing. Other important information such as that of the patient's medical history, demographics, medications, cardiovascular disease risk factors were obtained at the time the tests were done by the staff. The studies done in this paper used the conventional statistical techniques to find the accurate medical outcomes of the said problems. The techniques in the paper were LogitBoost, Artificial Neural Network, Locally weighted Nave Bayes, Bayesian Network, Support Vector Machine, Random Forests. As the date collected is highly imbalanced, SMOTE analysis was done for all the methods given. This paper presents the performance of Neural Networks using the gradient descent back-propagation algorithm using hidden units and the momentum thus evaluated using 10-fold cross validation. Bayesian Network classifier using Simulated Annealing algorithm achieves the highest Area Under Curve(AUC) value of 0.70. The results show that the performance of the RTF and SVM models using SMOTE has shown great improvement. The RTF and SVM achieve AUC of 0.91 and 0.71 respectively using the sampled dataset with 30% synthetic examples created in comparison to 0.9 and 0.57 respectively using the dataset without sampling. LogitBoost and Artificial Neural Network models have shown a slight improvement using by achieving Area under Curve of 69% and 63 % respectively without sampling and AUC of 70% and 67% with 0.3 created synthetic examples. From the different evaluation methods, the Random Tree Forest model was found to have achieved the highest accuracy using the 10-fold cross-validation method of 93%,using the holdout method 70/30 it was found to be 83% and holdout method 80/20 it turned out to be 88%.

The paper by [7] presents a very natural method of measuring of hypertension namely by tracking the activities done by any person. The lifestyle of the people were studied and it was found that people who had attended 4 or more years of college were better stress managers as this lead to more stress during the early stages of their life when they more mentally fit and hence it let them have better control over themselves. Social support also plays a vital role in stress management and it was found that those who had some kind of social support were better managers of stress. Living situations, financial difficulties were also some of the reasons for these issues. In this study, the results revealed that the regression model is statistically significant in predicting changes in HRQOL according to lifestyle-based intervention, stress, and social support. Educational levels, race, stress at baseline are significant predictors for predicting change in stress; social support at baseline is the significant predictor for predicting change in social support.

## III. DISCUSSION

Here we see that in [1] a system has been built which continuously learns about the different ways in which hypertension is happening and helps practitioners and clinicans as to how they can understand the early signs of hypertension and proper medication can be provided. In the [2] we see that the data we got has high prediction. It was better than [1] as here ANN was used and provided better accuracy. In [3] a new aspect was introduced of finding the height to waist ratio which was also found significantly important.[4] use pulse wave velocity to understand how can they me matched with heart rate. Here, they used data from 2007 and 2011 and tried to predict the possibility of hypertension with and without PWV. In the [5] paper, a new method of research was found using the cranial tissue of the brain and how can it be used to predict hypertension. Here they used Random Forest classifiers for the same. In [6] comparison have been done on the classification can be done using 6 different methods of comparing and contrasting. The paper shows that there are many different ways of comparing the different methods of predicting hypertension.

## IV. CONCLUSIONS AND FUTURE WORK

Real time prediction based on the dataset can reduce the number of readmission to the hospital and save the cost for the future and better predictability for future disease. In this paper we proposed models are used to diagnose hypertension. Apart from assisting doctors as a support tool it can have wide range of applications. It can be used to help nurse practitioners. Patients can be segregated into the possible problems that they may face due to hypertension using such models. Many different models have been summarized in this paper and they all have been done on cutting edge technologies. The prediction factors have ranged from 80-90%. The different techniques used and experimented to predict have been thoroughly studied and discussed. The papers analysed show that there is great research that has been done the different methods that hypertension can be related to such as waist to height ratio, lifestyle patterns etc. This shows that the different areas that have been researched. This adds that to the previous point that hypertension is a great problem in the world and hence needs to be addressed. This problem is set to increase with time and hence it is imminent that such vast amount of research is done on this topic. In the future work of the paper, there has been no mention of the psychological effects that lead to hypertension. Using the advanced brain mapping techniques available, the different emotions felt by people can be felt and they can also be used to find the cause of hypertension and its prediction. By mapping the brain waves with the actions done by the person on any day we can narrow down to exact amount of work done and which ones the person has the maximum stress doing. This shows what work gives him the most problem or tension leading to higher order of stress levels leading to other diseases.

References are important to the reader; therefore, each citation must be complete and correct. If at all possible,

references should be commonly available publications.

## REFERENCES


[1] ] Gitarja Sand et al.(2016) Health Risk Prediction for Treatment of Hypertension 2016 4th International Conference on Cyber and IT Service Management

[2] ] Daniel Lafreniere et al. (2016) Using Machine Learning to Predict Hypertension from a Clinical Dataset 2016 IEEE Symposium Series on Computational Intelligence (SSCI)

[3] Jung Ran Choi et al.(2018) Waist-to-height ratio index for predicting incidences of hypertension: the ARIRANG study BMC Public Health

[4] ] Teemu Koivistoinen et al. (2018) Pulse Wave Velocity Predicts the Progression of Blood Pressure and Development of Hypertension in Young Adults

[5] ] Federico Wadehn et al. (2017) Non-Invasive Detection of Intracranial Hypertension Using Random Forest Computing in Cardiology 2017; VOL 44 DOI:10.22489/CinC.2017.323-004

[6] Sharif et al. (2018) Using machine learning on cardiorespiratory fitness data for predicting hypertension Vol. 13, Iss. 4: e0195344.

[7] Mei-Lan Chen et al. (2018) Effect of a Lifestyle-Based Intervention on Health Related Quality of Life in Older Adults with Hypertension Journal of Aging Research Volume 2018, https://doi.org/10.1155/2018/6059560

[8] M. Czosnyka and J. D. Pickard (2017), Monitoring and interpretation of intracranial pressure, Journal of Neurology, Neurosurgery & Psychiatry 2004;75:813-821.

[9] X. Zhang, J. E. Medow, B. J. Iskandar, F. Wang, M. Shokoueinejad, J. Koueik, and J. G. Webster, Invasive and noninvasive means of measuring intracranial pressure: a review, Physiological Measurement, vol. 38, no. 7, pp. R143R182, 2017.

[10] C. Robba, D. Cardim, T. Tajsic, J. Pietersen, M. Bulman, J. Donnelly, A. Lavinio, A. Gupta, D. K. Menon, P. J. Hutchinson, et al., Ultrasound non-invasive measurement of intracranial pressure in neurointensive care: A prospective observational study, PLOS Medicine, vol. 14, no. 7, p. e1002356, 2017.

[11] Ashwell M, Gunn P, Gibson S. Waist-to-height ratio is a better screening tool than waist circumference and BMI for adult cardiometabolic risk factors: systematic review and meta-analysis. Obes Rev. 2012;13(3):27586.

[12] ] Pouliot MC, Despres JP, Lemieux S, Moorjani S, Bouchard C, Tremblay A, Nadeau A, Lupien PJ. Waist circumference and abdominal sagittal diameter: best simple anthropometric indexes of abdominal visceral adipose tissue accumulation and related cardiovascular risk in men and women. Am J Cardiol. 1994;73(7):4608.

[13] ] Kbi T, Khnen M, Iivainen T, Turjanmaa V. Simultaneous non-invasive assessment of arterial stiffness and haemodynamicsa validation study. Clin Physiol Funct Imaging. 2003;23:3136.

[14] U. Rimmele, R. Seiler, B. Marti, P. H. Wirtz, U. Ehlert, and M. Heinrichs, (e level of physical activity affects adrenal and cardiovascular reactivity to psychosocial stress, Psychoneuroendocrinology, vol. 34, no. 2, pp. 190198, 2009.