The Nucleus ***53****, No. 4 (2016) 236-242*

www.thenucleuspak.org.pk

The Nucleus

ISSN 0029-5698 (Print)

ISSN 2306-6539 (Online)

A Classification Approach Based on Genetic-Data-Structuring for the Prediction of Hypertension

A. Anis*, M.A. Fahiem and H. Tauseef

Department of Computer Science, Lahore College for Women University, Lahore, Pakistan

attiasultana22@gmail.com; itisab@gmail.com; humaiftikhar@hotmail.com

ARTICLE INFO

Article history:
Received : 09 September, 2016
Revised : 18 November, 2016
Accepted : 29 November, 2016

Keywords:
Hypertension
Genetic data
Alleles
Structured data
Naïve bays

ABSTRACT

Hypertension is known to be a major cause of death around the world. The death rate for this disease has increased up to 94% since last decade. Due to this disease, dangerous health conditions arise like heart failure, kidney failure and stroke. To prevent permanent loss, there is an imperious need for automated techniques to be developed for the detection of this disease. Recently, genetic information has been combined with machine learning techniques for the detection of hypertension disease. The genetics based diagnosis also plays a key role in evolution of this disease. In this paper both the genetic and demographic data is used forth detection of this disease. For detection purpose we have gathered genetic datasets of 250 patients from different databases (online repositories). Our research makes use of a feature set comprising genes, SNP alleles, risk alleles, chromosomes, region and nationality of patients. Here, an algorithm is proposed for data structuring. For performance evaluation of our proposed approach we have applied different types of classification models; naive bayes, filtered classifier, LWL, K and NBF network. The results obtained show that best accuracy is achieved using naive bayes i.e. 98%.*

1. Introduction

The major cause of death around the world is Hypertension (High Blood Pressure). Hypertension is a major risk factor for ischemic and haemorrhagic stroke, heart failure [1, 2], myocardial infarction, cognitive decline, chronic kidney disease and premature death. Poorly treated hypertension is usually associated with a progressive rise in blood pressure (BP). Hypertension is usually expressed with measurements, systolic blood pressure (SBP) and diastolic blood pressure (DBP). Normally hypertension exists when blood pressure exceeds 140/90 mm HG. There are two types of hypertension primary (essential) hypertension and secondary hypertension [4]. Primary hypertension gradually increases with age on the other hand secondary hypertension is normally based on various conditions e.g. serious sleep disorder [5], adrenal gland tumors, kidney problems, thyroid gland problem, certain medications, certain defects in blood vessels (inherited), decongestants (birth control pills, cold remedies, without counter pain relievers, prescription drugs and prohibited drugs).

With ageing factor high blood pressure normally exists [6]. Even in people under the age of 35, SBP exists. Even the risk factors that contribute to hypertension are inherited genetically. Genetics influence the inherited predisposition for certain diseases [7]. Genes' influence on hypertension has been suggested by family history that demonstrates associations of blood pressure among siblings and between parents and children [8]. Researchers are strongly concluding that genetic disorders and mutations ingenes can cause hypertension. Many other risk factors are also responsible like age, blood cholesterol level, blood pressure, diabetes, diet and stress that cause hypertension. Our research work is related to prediction of patient being hypertensive or not. For prediction our research work focuses on genetic and demographic data, which is gathered from different online repositories like dbSNP [9], Human Genome Mutation Database (HGMD) [10], 1000 Genome Project [11], GenBank [12], Genome-Wide Association Studies (GWAS) [13] and National Human Genome Research Institute (NHGRI) [14].

Six Sections are presented in this research; Section2 presents literature related to genetic and demographic data analysis based approaches. An improved frame work of proposed approach is discussed in Section 3, Section 4 presents classification, experimental results, conclusion and future recommendations are discussed in Section 5 and 6, respectively.

* Corresponding author

Table 1: Comparison of different online data repositories

Databases	No of features	Format	Type of data		No of records	Primary purpose
			Genetic	Demographic		
dbSNP [9]	11	ASN.1, XML	✓	✓	184 million	For biological researchers
HGMD [10]	13	ASN.1	✓	✓	About 2 million	Medical
1000 Genome Project [11]	8	BAM CRAM	✓	✓	3 billion	Detail catalogue for research
Gen Bank [12]	9	ASN.1,XML	✓	–	162million sequences	Health purpose
GWAS [13]	12	XML	✓	–	4000 SNP association	Medical
NHGRI [14]	11	ASN.1	✓	✓	About 2 million	For scientific research

2. Literature Review

There are many existing approaches for predicting hypertension, some researchers are focusing on genetic data and some are focusing on demographic data like age, weight, region, nationality etc.

Genetic data along with demographic data is available on different online repositories e.g. dbSNP [9], HGMD [10], 1000 Genome Project [11], GenBanK [12], GWAS [13], NHGRI [14], etc. A comprehensive comparison of these databases is given in Table 1.

2.1 Demographic Data Analysis Based Approaches

Different techniques have been used for predicting hypertension by using demographic data. Even data mining techniques have been used to examine, investigate and extract medical data by using difficult algorithms for disease prediction and unknown patterns. Oracle data miner tool has also used by researchers for predicting hypertension. This tool is used to investigate five parameters; diet, drug, weight reduction, smoke cessation and exercise for prediction [16]. Artificial neural networks (ANN) and fuzzy systems are used for predicting hypertension in patients using risk factors e.g. age, body mass index (BMI),SBP, DBP, smoking, obesity [17–19]. ANN, logistic regression and linear discrimination have also been used for prediction based on some variables like age, alcohol, weight, stress, obesity, lack of exercise, genetic factors and consumption of salt [20, 21]. Fuzzy expert systems introduced by some researchers are used for diagnosis of hypertension based on laboratory data like age, life style and heart rate. The linguistic variables, their values and the diagnosis process were modeled based on expert's information and from accessible literature [22, 23]. This disease is also diagnosed using bayesian statistics. Hypertension diagnosis has been done using compound pattern recognition method based on general, biochemical and blood pressure information [24]. Researchers have used four types of compound pattern recognition algorithms: two-stage classifier, random forest, and random subspace 3 and feature driven space division. The results of this approach are near to satisfactory [25]. Some researchers have used a two stage classifier for diagnosis of hypertension, first step is used to find essential or secondary hypertension and in second step human experts rules are used for making its own decision [26].

2.2 Genetic Data Analysis Based Approaches

The basis of hypertension is not well defined. However, the genes play an important role in the progression of this disease that cannot be ignored. Through genetic information hypertension disease can be diagnosed successfully. Various genes mutations are found, in some cases it may be genetic or it may be non-genetic. A gene is a segment of DNA that encodes function. Chromosome consists of a long strand of DNA which contains many genes.

The Bonferroni method is used to adjust multiple testing analyses. Single nucleotide polymorphism (SNP) genotype information is selected for diagnosis of diseases [27]. Multivariate logistic regression analysis of SNPs, associated with hypertension of different gender, male and female has been conducted. 33 SNPs of the 27 genes are selected on the basis of platelet, leukocyte, and other metabolic factors [28]. 3D scanning data association rule algorithm (ARA) and genetic algorithm (GA) are also being used for predicting hypertension [29]. SNP data is used for prediction of hypertension. For longitudinal binary outcomes two popular models, a marginal model and conditional model are used. For decision or predictions in individuals the conditional model performs better [30]. Recent studies have identified association between specific location, hypertension and high blood pressure. The aims of studies show that the genetic risk score (GRS) and SNPs are associated with hypertension. Longitudinal data of Korean population is used for predicting incidental hypertension. Unweighted GRS (cGRS) and weighted GRS (wGRS) are constructed from 4 SNPs related to high BP and hypertension. Cross-Sectional analysis reveals that cGRS are associated with

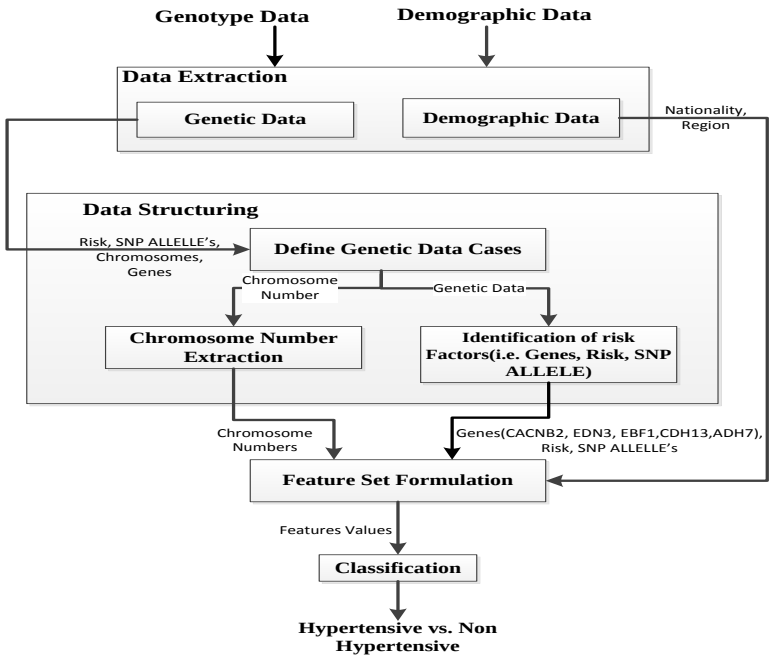

Fig. 1: Block diagram of proposed approach

prevalent hypertension and longitudinal analysis revealed that wGRS is associated with a greater risk of incident (unpleasant) hypertension. The data of risk variables (age, gender, present smoking status, SBP, family history of hypertension, BMI) and one genetic variable (cGRS or wGRS) derived from the 4 SNPs is used for constructing models. Finally, validity of the model was evaluated by k fold cross validation technique to obtain an overall result of 95% .

In one of the researches, statistical results are obtained for volunteer men and women of age group 30 to 65 years who are either drug patients for 1year or newly diagnosed patients [31]. To predict the risk of essential hypertension due to genes, individuals are genotyped for C825T genetic polymorphism of G protein β3 gene rs5443 by using restriction fragment length polymorphism [26].

ASNP is a DNA sequence variation occurring commonly within a population. These single nucleotides are A, T, C or Gin the genome. All common SNPs have only two alleles. Some authors have done their work with both type of data like age, gender cholesterol, glucose level, SBP, DBP and DNA damage data.

For this research, currently we are working on selected genes i.e. CACNB2, CDH13, EBF1, ADH7 and EDN3. These genes are involved in onset of hypertension disease.

3. Proposed Approach

This section is divided according to main processes; data extraction, data structuring, feature set formulation, classification and performance evaluation. Flow of our proposed approach is shown in Fig. 1.

3.1 Data Extraction

For research purposes we gathered data sets of patients from different databases. For this research we are considering both genetic and demographic data. Whenever data is gathered from different sources it is considered to be unstructured data. The extracted data is then used for further processing.

3.2. Data Structuring

Extracted data from different sources considered unstructured data. It is very tricky to extract "training datasets" from unstructured data. An algorithm is designed for structuring it. Algorithm is shown in Fig. 2.

This algorithm has different cases; in each case for selected genes it is checked whether the gene data consists of SNP allele (SA), Risk Allele (RA), location of genes in chromosomes, their region and nationality. Further in case one, gene CACNB2 will be given as an input if RA, SA and chromosomes exists related to this gene then it will be extracted and so on. All cases will be checked in a similar manner and will be inserted in tabular form. Sample dataset is presented below in Table 2. Furthermore, if any genes data do not have concise and complete information about patient like its RA, SA, chromosomes, its nationality and region, it will not be considered as a proper information of the patient and the data will be discarded.

Table 2: Structured data after execution of proposed algorithm

Patient	Genes					SNP Allele	Risk Allele	Chromosomes	Nationality	Region	Mutation
	CACNB2	EDN3	EBF1	CDH13	ADH7						
P1	✓	✓		✓		A/G, T/C	A,A,C	4,7	Yoruba	Yoruba	Yes
P2		✓	✓	✓		C/T,A/G, G/T	A,T,T	4,6,12	Japanese	Japanese	Yes
P3	✓					A/G	A	11	Chinese	Chinese	No
P4	✓	✓		✓		A/G,C/T, A/T	A,T,T	8,15,19	Japanese	Japanese	Yes
P5	✓		✓		✓	C/T,A/G, A/T	T,T,C	6,3,11	European	European	Yes
P6			✓			A/T	C	9	European	European	No

```
Id, n, Gen, RA, SA, Chr, Region, N, d [m, j] = 0
// RA= Risk Allele, SA= SNP Allele.
// Chr= Chromosomes, Gen= Gene.
For m (1 to 6)
For j (1 to n)
Switch (Gen) {
//Gene match with RA, SA, Chromosomes in all
 Cases.
Case 1:
Do If "Gen exists in chromosomes".
        If "CACNB2 exist in chromosomes
        Then m [1, 2] =RA
m [1,3] = SA
else break;
Case 2:
do If "Gen exists in chromosomes".
        If "EBF1 exists in chromosomes
        Then m [2, 2] =RA
m [2,3] = SA
else break;
Case 3:
do If "Gen exists in chromosomes".
        If "EDN3 exists in chromosomes
        Then m [2, 2] = RA
m [2,3] = SA else break;
Case 4:
do If "Gen exists in chromosomes".
        If "CDH13 exist in chromosomes
        Then m [4, 2] =RA
m [4,3] = SA
else break;
Case 5:
do If "Gen exists in chromosomes".
        If "ADH7 exists in chromosomes
        Then m [5, 2] = RA
m [5,3] = SA
else break;   end switch
}
```

Fig. 2: Algorithm for inserting data in table

3.3 Feature Set Formulation

A proper feature set with genes, SA, RA, chromosomes, nationality and region is extracted for further processing after structuring. SA is a variation in a single nucleotide which may occur at some specific position in the genome and RA sometimes influence alleles which shows the onset or severity of the disease in affected people.

4. Classification

For prediction we have used different classifiers. The complete feature set based on seven features are fed to different classifiers. Classifiers used for prediction are discussed below. Naive bayes is highly scalable and simple probabilistic classifier used for independence assumption about features. Strong results are obtained using naive bayes classifier.

4.1 Locally Weighted Learning (LWL)

For assigning weights to an instance, LWL uses an instance-based algorithm which is then used by a specified weighted instances handler. The specified classifier does not work if any column is marked as incompatible. Sometimes error occurs that no nominal values are available.

4.2 K-Star (K*)

This is an instance-based classifier, which contains the test instance that is based upon the class of those training instances similar to it, as determined by some similarity function. It uses an entropy-based distance function to differ the other instance-based learners. Hypothesis is directly constructed from the training instances so it is called instance -based.

4.3 Naive Bayes

Naïve bayes takes linear time for constructing classifier. It is a simple technique. The given instance of a problem to be classified, represented by a vector $x=(x_1 \ldots\ldots x_n)$, n features are independent variables. It assigns probabilities to this instance

$$p(Ck|x1, \ldots\ldots\ldots xn) \quad (1)$$

$$Class(v) = argmaxp(Ck) \prod_{i=1}^{n} \quad P\left(\frac{\aleph i}{Ck}\right) \text{k} \epsilon \{1,\text{k}\} (2)$$

Where P ($\mathbf{C_k}$) is the probability of $class_i$ given a sample candidate region using classifier k and a_k is the weight associated to the probabilistic prediction of class $\mathbf{C_k}$. It gives more accuracy rate than other classifiers due to its sophisticated classification methods.

4.4 Filtered Classifier

The filter structure is based completely on the training data and the filter processes the test instance without changing their structure. But in our case this classification method is not useful for prediction and classification.

4.5 Radial Basis Function (RBF) Network

RBF network is a popular machine learning algorithm. It is used for classification. It has three layers. It quickly trains data but after training it runs slowly. It does not maintain large numbers of training samples or large numbers of features in the input space well. A prediction accuracy rate of 50% is achieved when experimenting proposed approach.

5. Experimental Results

The results of all classifiers are illustrated in Table 3. The Naive Bayes yields more powerful accuracy rate i.e. 0.981 because of its strong classification techiques. Weka tool is used for experimentation and prediction.Patient data is gathered form different sources and given as a input.The accuracy rate of naïve bayes is higher than other classifiers. In naive bayes, the probabilities for each attribute are calculated independently from the training dataset. It needs enough data to understand the probabilistic relationship of each attribute in isolation with the output variable. It produces good results, like highest accuracy rate which implies the number of people predicted as a hypertensive patient. Sensitivity shows true positive rate and specificity shows true negative rate. The overall system performance is satisfactory as is evident from Fig. 3.

Table 3: Performance of different classifiers using proposed approach

Classifier	Sensitivity	Specificity	Accuracy
Naive Bayes	0.93	0.94	0.981
Filtered Classifier	0.83	0.84	0.841
LWL	0.82	0.82	0.821
K*	0.73	0.74	0.751
RBF Network	0.74	0.67	0.73

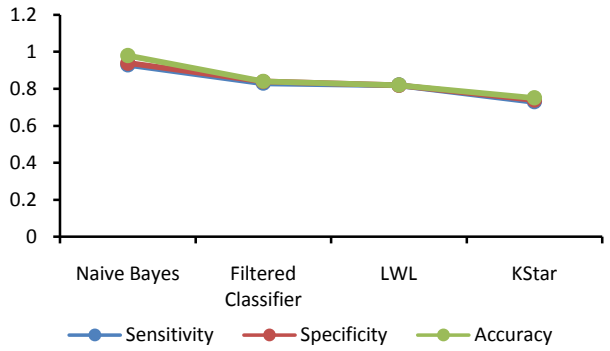

Fig. 3: Graphical representation of results

6. Conclusion and Future Recommendation

Hypertension is a global disease that is considered to be a major cause of death all over the world. In this research, we have introduced a new approach for predicting this disease. In proposed research we have gathered genetic and demographic data from different sources for structuring using the proposed algorithm and then prediction of Hypertension is done. Different classifiers are applied on structured data for prediction. In this research currently we are working on some selected genes of patients. In future large group of genes, complete genetic information and more features can be added for prediction. For prediction purpose WEKA tool is used. Naive Bayes has shown better results of prediction. The results are shown in Table 3. Comparison of existing approaches with our approach is given in Table 4. The accuracy rate of proposed system is 98%, it is much better than previous techniques. Actually we have focused on genetic data comprising of genes, alleles and chromosome position of different patients. A datasets of 250 patients is used for prediction. Naive Bayes produced highest accuracy. It is more efficient for making decision and prediction of disease even if patient's information about genes is incomplete, not present or gene related information cannot be extracted. Naive Bayes can even give proper results in cases with minimum information.

Table 4: Comparison of proposed approach with existing techniques

Author	Risk factor	Tools and Techniques	Accuracy
Aljumah et al. [15]	Diet, drug smoke cessation, weight reduction, exercise	Oracle data miner tool	84%
Kaur, et al. [16]	Age, BMI, SBP, DBP	Artificial intelligence, Fuzzy system	82%
Djam et al. [18]	Age, BMI, SBP, DBP, alcohol assumption.	Artificial intelligence, Fuzzy system	81%
Shehu et al. [19]	Age, alcohol, weight, stress	Neural networks, logistic regression, linear discrimination.	94%
Abrishami et al. [20]	Obesity, Lack of exercise, Genetic factor, Consumption of salt.	Neural networks, logistic regression, linear discrimination.	89%
Abdullah et al. [21]	Age, life style and heart rate.	Fuzzy experts system	95%
Das et al. [22]	Age, life style and heart rate.	Fuzzy experts system	90%
Blinowska et al. [23]	Blood pressure, clinical and biochemical information	Bayesian statistics	93%
Krawczyk et al. [24]	General, biochemical and blood pressure information	Compound pattern recognition method	96%
Woznaik.et al. [25]	Essential hypertension, secondary hypertension, blood pressure.	Two stage classifier	92%
Izawa et al. [27]	SNPs of Genes	Multivariate logistic regression analysis of SNPs	88%
Chiu et al. [28]	3-D scanning data	ARA, genetic algorithm	89%
Choi et al. [29].	Genetic data	Marginal data, conditional data	92%
Lim et al. [30]	Age, gender, SBP, smoking status, family history of hypertension, BMI	Multiple linear regression, logistic regression analysis	94%
Hemimi et al. [31]	Genetic data (C825T genetic polymorphism of G protein β3 gene)	Restriction fragment length polymorphism (statistical analysis)	92%
Proposed Approach	Genetic and demographic data	Naïve Bayes	98%

References

[1] D. Gasecki, M. Kwarciany, K. Kowalczyk, A. Rojek , T. Nowicki, M. Skrzypek-Czerko, E. Szurowska, P. Boutouyrie, S. Laurent and K. Narkiewicz, "Aortic stiffness is an independent biomarker of subclinical brain damage in acute ischemic stroke", J. Hypertension, 33(Supp 1): pp. e1-e524, June 2015.

[2] N. Hussain, D. Capener, C. Elliot, R. Condliffe, J. M. Wild, D. G. Kiely and A. Swift. "Interventricular septal angle can be used to predict which patients have combined post capillary or pre capillary pulmonary hypertension in left heart disease", J. Cardiovascular Magnetic Resonance, 17, Supp. 1, p. 338,Feb 2015.

[3] Y. Yano, J. Stamler, D.B. Garside, M.L. Daviglus, S.S. Franklin, M.R. Carnethon, K. Liu, P. Greenland and D. M. Lloyd-Jones, "Isolated systolic hypertension in young and middle-aged adults and 31-year risk for cardiovascular mortality", J. American College of the Cardiology, vol. 65, no. 4, pp. 327-335, Feb 2015.

[4] K. Chau, D. Holmes, A. Melck and C. Chan- Yan, "Secondary hypertension due to mutant aldosterone-producing adenoma and parathyroid adenoma", American Journal of Hypertension, vol. 28, no. 2, pp. 280-282, Feb 2015.

[5] A.P. Steinhorst, S.C. Gonçalves, A.T. Oliveira, D. Massierer, M. Gus, S.C. Fuchs, L.B. Moreira, D. Martinez and F.D. Fuchs, "Influence of sleep apnea severity on blood pressure variability of patients with hypertension, "Sleep and Breating, vol. 18, no. 2, pp. 397-401, May 2014.

[6] A. Harvey, A. C. Montezano and R.M. Touyz, "Vascular biology of ageing- Implications in hypertension, "Journal of Molecular and Cellular Cardiology, vol. 83, pp. 112–121, June 2015.

[7] X. Xu, S. Su, F. A. Treiber, R. Vlietinck, R. Fagard, C. Derom, M. Gielen, J.F. Ruth. Loos, H. Snieder and X. Wang, "Specific genetic influences on night time blood pressure", American Journal of Hypertension, vol. 28, no.4, pp. 440-443, April 2015.

[8] J.L. Greaney, E.L. Matthews and M.M. Wenner, "Sympathetic reactivity in young women with a family history of hypertension", American Journal of Physiology-Heart and Circulatory Physiology**,** vol. 308, no. 8, pp. H816-H22, April 2015.

[9] A. Kitts, L. Phan, M. Ward and J. B. Holmes, "The database of short genetic variation (dbSNP)", [ONLINE], Available: http://www.ncbi.nlm.nih.gov/.

[10] P.D. Stenson, M. Mort, E. V. Ball, K. Howells, A. D. Phillips, N.S. Thomas and D.N. Cooper, "The human genome mutation database: 2008 update", Genome Medicine, vol. 1, no. 1, pp. 1, Jan. 2009.

[11] EMBL-EBI, "1000 genome", The International genome sample resource", European 2016, [ONLINE] http://browser.1000 genomes.org/index.html [Accessed August 6, 2016].

[12] "NIH Genetic Sequence Database", Japan 2013, [ONLINE] https://www.ncbi.nlm.nih.gov/genbank/ [Accessed June 7, 2016].

[13] "GenomeWide–Association Study". [ONLINE] https://www.ebi. ac.uk/gwas/ [Accessed, March 4, 2016].

[14] "Genome" [ONLINE], USA 2015, https://www.genome.gov/, [Accessed April 23, 2016].

[15] A.A. Aljumah, M.G. Ahamad and M.K. Siddiqui, "Predictive analysis on hypertension treatment using data mining approach in Saudi Arabia", Intelligent Information Management", vol. 3, No. 6, pp. 252-261, Nov. 2011.

[16] A. Kaur and A. Bhardwaj, “Artificial Intelligence in hypertension diagnosis: A review”, Int. J. Comp. Sci. Inf. Technol., vol. 5, no. 2, pp. 2633-2635, April 2014.

[17] R. Samant and S. Rao, "Evaluation of artificial neural networks in prediction of essential hypertension", Int. J. Comp. Appl., vol. 81, no. 12, pp. 34-38, Nov. 2013.

[18] X.Y. Djam and Y.H. Kimbi."Fuzzy expert system or the management of hypertension", The Pacific Journal of Science and Technology, vol. 12, no.1, pp. 390-402, May 2011.

[19] N. Shehu, S.U. Gulumbe and H.M. Liman, "Comparative study between conventional statistical methods and neural networks in predicting hypertension status", Advances in Agriculture, Sciences and Engineering Research, vol. 3, no. 5, pp. 867-874, 2013.

[20] Z. Abrishami and H. Tabatabaee, “Design a fuzzy expert system and a multi-layer neural network system for diagnosis for hypertension”, MAGNT Research Report, vol. 4, no.11, pp. 138-145, October 2015.

[21] A.A. Abdullah, Z. Zakaria and N. F. Mohammad, "Design and development of fuzzy expert system for diagnosis of hypertension", Second International Conference on Intelligent Systems, Modelling and Simulation, Kuala Lumpur, Malaysia and Phnom Penh, Cambodia, pp. 113-117, January, 2011.

[22] S. Das, P.K. Ghosh and S. Kar, "Hypertension diagnosis: A comparative study using fuzzy expert system and neuro fuzzy system", Fuzzy Systems, IEEE Int. Conf., Hyderabad, India, pp. 1-7, July,2013.

[23] A. Blinowska, G. Chatellier, J. Bernier and M. Lavril, “Bayesian Statistics as applied to hypertension diagnosis”, IEEE Transactions on Biomedical Engineering, vol. 38, no. 7, pp. 699-706, July1991.

[24] B. Krawczyk and M. Wozniak, “Hypertension diagnosis using compound pattern recognition method”, J. Med. Inf. Technol., vol.18, pp. 41-50, 2011.

[25] M. Wozniak, “Two-Stage classifier for diagnosis of hypertension type”, Lecture Notes in Bioinformatics, Springer Berlin Heidelberg, vol. 4345, pp.433-440, 2006.

[26] X. Yang, J. He, D. Gu, J.E. Hixson, J. Huang, D.C.Rao, L.C. Shimmin, J. Chen, T. K. Rice, J. Li and K.S. Kelly, “Associations of epithelial sodium channel genes with blood pressure changes and hypertension incidence: The Gen Salt Study”, American Journal of Hypertension, ISSN1941-7225, pp. hpu060, April 2014.

[27] H. Izawa, Y. Yamada, T. Okada, M. Tanaka, H. Hirayama and M. Yokota, "Prediction of genetic risk for hypertension", Hypertension, vol.41, no.5, pp.1035-1040, May2003.

[28] C. Chiu, K.H. Hsu, P.L. Hsu, C.I. Hsu, P.C. Lee, W.K. Chiou, T.H. Liu, Y.C. Chuang and C.J. Hwang, "Mining three-dimensional anthropometric body surface scanning data for hypertension detection", IEEE Transactions on Information Technology in Biomedicine, vol.11, no.3, pp.264-273, May2007.

[29] Y.H. Choi, R. Chowdhury and B. Swaminathan, “Prediction of hypertension based on the genetic analysis of longitudinal phenotypes: A comparison of different modeling approaches for the binary trait of hypertension”, BMC Proc., vol. 8, no. 1, June 2014.

[30] N. K. Lim, J. Y. Lee, J.Y. Lee, H.Y. Park and M. C. Cho, “The role of genetic risk score in predicting the risk of hypertension in the Korean population: Korean genome and epidemiology study”, PloSOne, vol. 10, no.6, June 2015.

[31] N. Hemimi, A.A. Mansour and M.M. Abdelsalam, “Prediction of the risk for essential hypertension among carriers of c825t genetic polymorphism of g protein β3(GNB3) gene”, Biomarker Insights, vol. 11, pp.69–75, May 2016.