

An Optimized DBN-based Coronary Heart Disease Risk Prediction

K. Lim, B.M. Lee, U. Kang, Y. Lee

Kahyun Lim^{1*}, Byung Mun Lee², Ungu Kang², Youngho Lee²

1. Department of IT Convergence Engineering; 2. Department of Computer Engineering
Gachon University, Republic of Korea

1342 Seongnamdaero, Sujeong-gu, Seongnam-si, Gyeonggi-do, Korea

bmlee@gachon.ac.kr, ugkang@gachon.ac.kr, lyh@gachon.ac.kr

*Corresponding author: lyh@gachon.ac.kr

Abstract: Coronary Heart Disease (CHD) is the world's leading cause of death according to a World Health Organization (WHO) report. Despite the evolution of modern medical technology, the mortality rate of CHD has increased. Nevertheless, patients often do not realize they have CHD until their condition is serious due to the complexity, high cost, and the side effects of the diagnosis process. Thus, research on predicting CHD risk has been conducted. The Framingham study is a widely-accepted study in this field. However, one of its limitations is its overestimation of risk, which threatens its accuracy. Therefore, this study suggests a more advanced CHD risk prediction algorithm based on Optimized-DBN (Deep Belief Network). Optimized-DBN is an algorithm to improve performance by overcoming the limitations of the existing DBN. DBN does not have the global optimum values for number of layers and nodes, which affects research results. We overcame this limitation by combining with a genetic algorithm. The result of genetic algorithm for deriving the number of layers and nodes of Optimized-DBN for CHD prediction was 2 layers, 5 and 7 nodes to each layers. The accuracy of the CHD prediction algorithm based on Optimized-DBN which is developed by applying results of genetic algorithm was 0.8924, which is better than Framingham's 0.5015 and DBN's 0.7507. In the case of specificity, Optimized-DBN based CHD prediction was 0.7440, which was slightly lower than 0.8208 of existing DBN, but better than Framingham's 0.65. In the case of sensitivity, Optimized-DBN is 0.8549, which is better than Framingham 0.4429 and DBN 0.7468. AUC of suggesting algorithm was 0.762, which was much better than Framingham 0.547 and DBN 0.570.

Keywords: Artificial Neural Networks (ANN), Deep Belief Network (DBN), Coronary Heart Disease (CHD), computational intelligence, genetic algorithm, CHD prediction.

1 Introduction

Coronary Heart Disease (CHD) is a disease in which a waxy substance called plaque builds up inside the coronary arteries [11]. It has been a leading cause of death globally for 15 years according to a World Health Organization report [1]. Once CHD is developed, it is nearly impossible to cure completely and patients are at risk of sudden death due to sudden myocardial infarction or etc. Therefore, the accurate risk prediction of CHD is very important in that it gives patients a chance to live better life by caring themselves more cautiously [6].

With this reason, there have been many researches to develop the method for predicting the risk of CHD. Among these, the Framingham Heart Study which was developed through prospective cohort study and the US Adult Treatment Panel (ATP) are widely used as standards when predicting the risk of CHD [15]. And as Artificial Neural Networks (ANN) are becoming increasingly popular because of its outstanding classifying performance, it is begin to be adopted into establishing the risk prediction model for CHD [20]. Research on developing a CHD risk

prediction model based on various ANN algorithms such as Random Forest (RF), Support Vector Machine (SVM), and Deep Belief Network (DBN) have been actively conducted. Although the risk prediction model based on ANN algorithms boasts a much better performance than the previous model, nevertheless, there is still room for improvement [9]. Therefore, these days, to develop a more improved version of prediction model, an ensemble technique which means combination of two different algorithms has been attempted [21]. The performance depends on how to combine two algorithms and if those are able to complement each other's disadvantages, it can significantly improve the prediction performance. However, developing CHD risk prediction model by combining two ANN algorithms have not tried much yet.

Among many ANN algorithms, Deep Belief Network is one of the most widely-used ANN algorithm because it has advantage in overfitting issue compared to other ANN algorithms [5]. Overfitting is one of the bothersome issue to most of the ANN algorithm researchers because it degrades prediction performance when applied to actual field [14]. Through regulating the parameter values of hidden or visible layers, DBN can be able to solve the overfitting problem to a certain degree. Though, DBN also has a limitation that there is no optimal value adequate to all datasets. The optimal values of each parameters are differed depending on datasets. Thus, researchers set values randomly and try several times, and select the values which show the best performance. However, there is no conviction that it is the optimal value set.

Therefore, in this study, our goal was to develop an CHD risk prediction model based on Optimized-DBN by combining DBN with Genetic Algorithm (GA). GA is an optimization algorithm which is useful for finding the optimal value. We utilized the GA to find the optimal value of the DBN parameter values of hidden and visible layers which is adequate to the dataset to predict the risk of CHD.

In section 2, we explained background knowledge about DBN and GA which are needed to understand the Optimized-DBN algorithm we suggest through this paper. In section 3, we described process and methodology of research to develop and verifying Optimized-DBN. The results of research such as performance comparison with the traditional DBN and other ANN algorithms which are widely being used now are recorded in section 4. Lastly, we concluded the paper with suggestion of future works in section 5.

2 Backgrounds

2.1 Deep belief network (DBN)

Deep learning refers to a collection of algorithms that pursues the improvement of performance by setting multiple hidden layers and passing information to those hidden layers several times. In 2006 Jeffrey Hinton, a professor at the University of Toronto who is called the master of artificial intelligence, first proposed the DBN in the paper "A fast learning algorithm for deep belief nets" [7]. With this paper, he proved that a DBN can greatly improve existing artificial neural networks, which take a lot of time to optimize and perform worse through the unsupervised pre-learning method. A DBN consists of a restricted Boltzmann machine (RBM) and artificial neural network and the RBM is composed of visible neurons and hidden neurons. Once the RBM obtains the training result of the input data, it uses the result as the input data for the next layer. It accumulates hidden neurons and learns [16]. The energy function of RBM is as in the following equation (1).

$$E(V, H) = - \sum_{i,j} v_i h_j - \sum_{i \in v} a_i v_i - \sum_{j \in H} b_j h_j \quad (1)$$

In equation (1), v_i and v_j are the binary values of the visible layer of i and the hidden layer

of j . In addition, h_i and h_j mean the bias of the visible layer of i and the hidden layer of j . W means a weighting matrix. Therefore, the energy function of the RBM and the Boltzmann machine without bias is the same as the following equation (2).

$$E(V, H) = -\frac{1}{2} \sum_{i,j} w_{ij} v_i h_j = -\frac{1}{2} V^T W H \quad (2)$$

In this case, the network V and H vectors follow the probability of the following equation (3).

$$p(V, H) = \frac{1}{Z} e^{-E(V, H)} \quad (3)$$

Z is the partition function, and p satisfies a probability value between 0 and 1. The updating of the weights uses k -step contrastive divergence (CD- k). $\langle v_i h_j \rangle$ data is the mean value of the input patterns of the visible layer of i and the hidden layer of j , and $\langle v_i h_j \rangle$ model represents the average value of the network [3].

The DBN sets the initial weights by performing unsupervised pre-training before performing the error-propagation algorithm [9]. After the initial weight value is determined, it performs an error back-propagation algorithm with the supervised learning method and performs fine-tuning to optimized performance [22]. However, research on the optimum value of the number of layers and nodes is still lacking, especially when performing pre-training through RBM.

The number of layers and nodes is a variable that influences research results, but there is a limit to the global optimal value because the optimal value depends on the type and characteristics of the dataset to be learned. Most studies using DBN have been conducted by researchers to set several values randomly and apply values which showed the best performance. However, this method has limitations because it is difficult to confirm whether its result is the optimum result. Our study overcomes this problem of DBN by applying a genetic algorithm.

A genetic algorithm (GA) is a method to find the optimum value. GA simulates a mathematical evolutionary method using crossover, mutation and selection. In other words, GA can be used to generate an optimal value of the number of layers and nodes of the DBN to maximize its performance. Through a genetic algorithm, we improve the performance of the CHD risk prediction by finding the optimal number of layers and nodes for the data through a genetic algorithm and applying it to the DBN.

2.2 Genetic algorithm (GA)

Genetic algorithms were created by John Holland [8], a professor of computer science and psychology at the University of Michigan. This is an algorithm in which the initial population of artificial chromosomes is reproduced through crossover and mutation, and in the process, chromosomes with a low fitness value are extinguished and chromosomes with a high fitness value survive. Survived chromosomes thus dominate the population. A genetic algorithm is an optimization algorithm performed in this form [13]. The genetic algorithm assigns the rank of each object through the fitness function and sets as many genes as the researcher set. It transmits the characteristics of genes having high fitness by conveying those to the next generation. After that, the next generation of chromosomes is generated through a crossover process. The crossover process has the effect of accelerating the convergence speed to its optimal value. Finally, through the mutation process which changes chromosomes at a certain probability, the global search effect is maximized in this stage, converging toward its optimum value. The genetic algorithm is based on the schema theorem. The schema is a set of strings consisting of 0, 1, and asterisks (*), and the genetic algorithm is executed by adjusting these schema. When the reproduction rate in the

genetic algorithm is proportional to the fitness, the probability that a particular schema (H) will survive the next generation can be predicted by the following equation (4) [19].

$$P_H^{(C)} = 1 - P_c \left(\frac{l_d}{l-1} \right) \quad (4)$$

The probability that schema H will survive after the mutation is as following equation (5). The genetic algorithm maximizes the global search effect by the mutation process.

$$P_H^{(m)} = (1 - P_m)^n \quad (5)$$

It is possible to get the probability of the schema growing according to the above two expressions, The formula is as following equation (6)

$$m_H(i+1) = \frac{\hat{f}_H(i)}{\hat{f}(i)} m_H(i) \left[1 - p_c \left(\frac{l_d}{l-1} \right) \right] (1 - p_m)^n \quad (6)$$

In this study, a genetic algorithm was used to derive the optimal DBN layer and node number values. The initial population size, N, was set to 256, P_c was set to 0.7, P_m was set to 0.001, and the termination condition was set to 200 generations following regular settings. The sigmoid function was also used as a fitness function.

3 Optimized DBN based coronary heart disease risk prediction

3.1 Experimental design

The purpose of this study was to develop and validate a predictive algorithm for coronary heart disease risk based on an optimized-DBN. The optimized-DBN improves performance by applying optimal values for the layer and node number derived by genetic algorithm to DBN. We developed a predictive algorithm for CHD risk through CHD risk factor data and verified it using a confusion matrix and ROC curve. The research method for this paper is shown in Figure 1. The training data was applied to the RBM, and when the RBM configuration was completed, the DBN was constructed through the error propagation algorithm. Then, the optimum value of the number of nodes and layers of the DBN was derived through the genetic algorithm, and the optimized result was applied to generate the optimized-DBN. The performance of the optimized-DBN based CHD risk prediction was verified by the confusion matrix and the ROC curve. And we utilized Matlab as a toolbox for developing ANN model and R 3.0 to visualize results.

3.2 Data

The optimized-DBN based CHD risk prediction proposed by this study predicts the risk CHD of patients through relevant variables. Data from the 6th National Health and Nutrition Examination Survey (KNHANES) was used for training and validation [12]. The KNHANES is a national health and nutrition survey conducted by the Korea Center for Disease Control and Prevention (KCDC) to identify the status of and trends in people's health and nutritional status [10]. This data includes the CHD risk factor data and CHD incidence data for each subject, which is useful for training and testing for the CHD risk prediction [17]. This study also utilized the latest data from the National Health and Nutrition Examination Survey of Korea. Variables for predicting the risk of CHD were as in the Framingham study. Eight factors including gender, age, systolic blood pressure (SBP), diastolic blood pressure (DBP), total cholesterol (TCL), high density cholesterol (HDL), obesity, and smoking were set as predictors of CHD. We defined

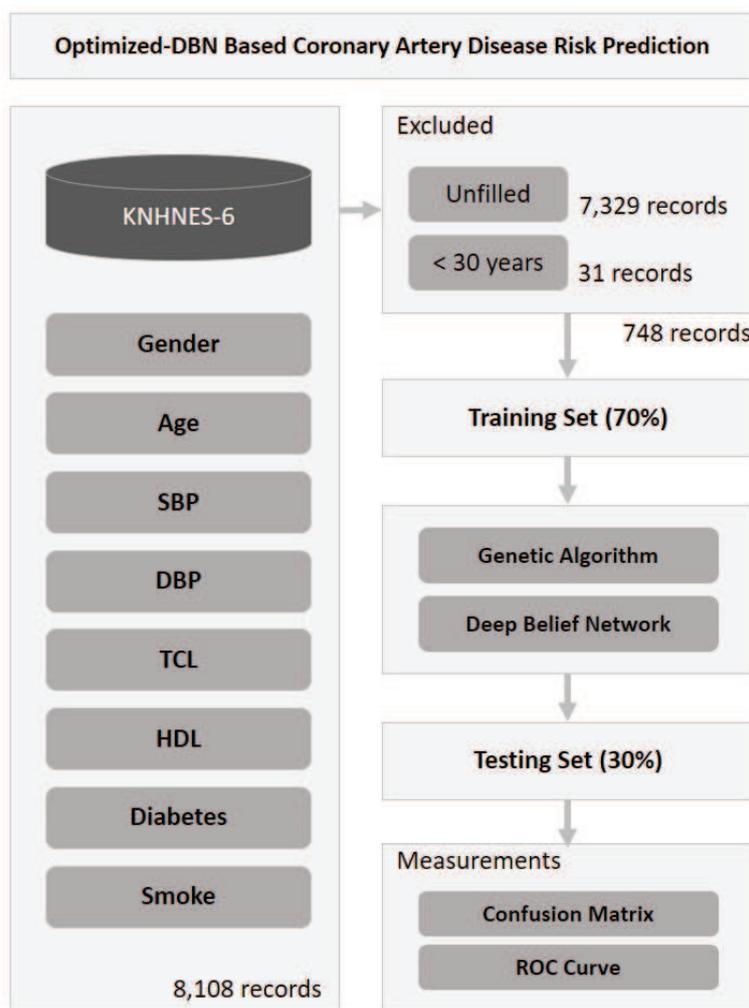


Figure 1: Experimental design for developing optimized-DBN based coronary heart disease risk prediction

age, total cholesterol (TCL), high density cholesterol (HDL), systolic blood pressure (SBP), and diastolic blood pressure (DBP) as continuous variables and gender, smoking, and obesity as nominal variables. The above 8 variables were extracted from KNHANES-6 data.

A total of 8,108 patients were enrolled in the KNHANES-6 study. 7,329 patients were excluded from the study because of the lack of data on risk factors for CHD suggested by the Framingham study or the presence of CHD. Furthermore, as in the Framingham study, data for subjects under the age of 30 was excluded. The remaining 748 data were used for analysis and prediction. In addition, 70% of the data was randomly extracted and used as training data for the creation of the optimized-DBN. The remaining 30% of the data was used to evaluate the performance of the optimized-DBN.

Of the KNHANES-6 data, 748 data points met the conditions of the study. Therefore, this data was used for research. A randomly chosen 70% of this training data was applied to the RBM first. The RBM uses the training result as training data for the next layer. In this way, the RBM is constructed by stacking hidden layers. Once the RBM configuration is complete, the DBN is configured via the error propagation algorithm. In this process, the sigmoid function is used as the activation function. A genetic algorithm is then performed to derive the optimum value of the number of nodes and layers of the DBN. The initial population size (N) of the

	Low Risk (288 patients)		High Risk (260 patients)	
	Mean	SD	Mean	SD
AGE	50.11	14.102	53.06	13.775
TCL	206.69	39.587	201.40	39.569
HDL	43.99	9.569	43.54	9.947
SBP	123.78	16.040	123.11	15.696
DBP	79.56	11.200	78.50	11.066
GENDER	M 302, W 186		M 135, W 107	
SMOKING	Y 318, N 170		Y 169, N 91	
DIABETES	Y 463, N 25		Y 196, N 64	

Figure 2: Summary of data from Korea National Health and Nutrition Examination Survey according to Framingham Risk Score Guidelines

genetic algorithm was set to 256, the possibility of crossover (P_c) was set to 0.7, the possibility of mutation (P_m) was set to 0.001, and the termination condition was set to 200 generations. After the genetic algorithm is implemented, the optimal value of the layer and node values with the lowest error rate is selected, and this value is applied to generate Optimized-DBN. When the error rate gradually decreases over the generations and converges toward its lowest value, the RBM performs learning for CHD risk prediction. As a result of the genetic algorithm, the number of layers was 2 and the number of nodes was 5 and 7 nodes to each layers, and this value was applied to the optimized-DBN configuration. After the optimized-DBN is constructed, an error backpropagation algorithm is performed, performance is improved through fine-tuning, and the optimized-DBN configuration is completed.

3.3 Performance measure

In this study, confusion matrix and receiver operating characteristic (ROC) curves were used as performance evaluation indexes. The confusion matrix can be used to assess the performance of the proposed algorithm by comparing the predicted risks with actual values. TP (true positive) means that an actual CHD risk patient is correctly predicted as a CHD risk patient. TN (true negative) means that patients with low CHD risk are correctly predicted as those with low CHD risk. These two indicators are values that are precisely classified [18]. FP (false positive) means that a person with a low CHD risk is misdiagnosed as a CHD risk patient. False negative (FN) means that patients with high CHD risk are incorrectly classified as those with low CHD risk. These two indicators are not well classified [2]. The ROC curve is a graph showing how sensitivity and specificity are related to each other. In this study, the ROC curve was used to show the accuracy of the optimized-DBN based CHD risk prediction. Sensitivity means the probability that the prediction is correct when the prediction algorithm assumes that the CHD risk is high. It is a measure of how accurately the proposed algorithm predicts patients with high CHD risk. Specificity means the probability that the prediction is correct when the prediction algorithm determines that the CHD risk for a patient is low. It is a measure of how accurately the proposed algorithm predicts patients with low CHD risk. The ROC curve can be interpreted

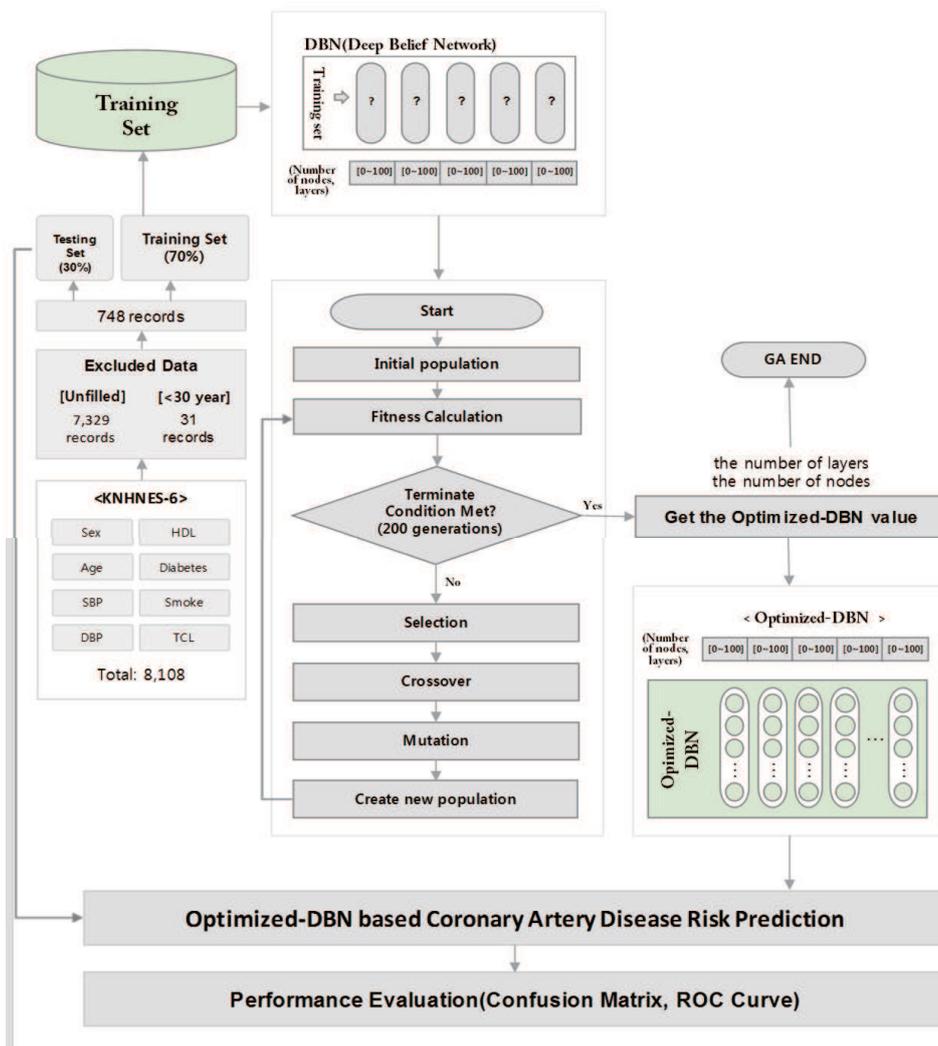


Figure 3: Process for developing optimized DBN based coronary heart disease risk predictions

as the proposed algorithm’s accuracy increasing as the graph converges to 1. In other words, as the area under the curve (AUC), which is the area under the line of the ROC curve graph, approaches 1, it can be estimated that the accuracy of the prediction algorithm is high.

4 Results

4.1 Genetic algorithm results

Our results show that the proposed CHD risk prediction algorithm performs better than the existing CHD risk prediction algorithm. The proposed CHD risk prediction algorithm is based on DBN, which is an artificial neural network algorithm. We overcome the limitations of DBN by using a genetic algorithm. After constructing the RBM through the training data and finishing the DBN configuration, the genetic algorithm was applied to derive the optimal value of the number of nodes and layers. The initial population (N) of the genetic algorithm was set to 256, the possibility of crossover (Pc) was set to 0.7, the possibility of mutation (Pm) was set to 0.001, and the termination condition was set to 200 generations. As shown in Figure 4, the error rate gradually decreased over the generations. The error rate, which was close to the initial 3.8,

decreased as the number of generations increased. It can be seen that the error rate converges to 0.3033 near the termination condition of 200 generations. After the genetic algorithm was terminated, the layer and node values with the lowest error rate were selected as the optimal solution. The value was 2 layers and 5 and 7 nodes to each layers. After applying these values to the DBN, the optimized-DBN generation was completed by supervised learning.

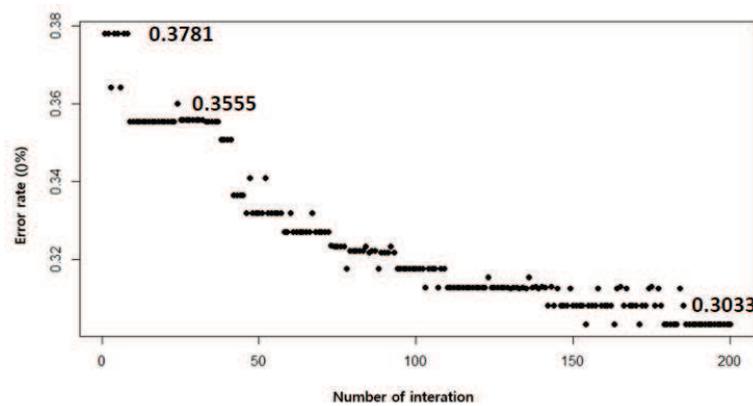


Figure 4: Error rate of genetic algorithm to develop optimized-DBN

<ul style="list-style-type: none"> • Sensitivity : 0.855 • Specificity : 0.744 • Accuracy : 0.829 		Actual result / Classification	
		High (1)	Low (0)
Predictive Result / Classification	High (1)	TP 837	FP 75
	Low (0)	FN 142	TN 218

Figure 5: Confusion Matrix of optimized-DBN based CHD risk prediction

The confusion matrix of the optimized-DBN based CHD risk prediction is shown in Figure 5. TP, which means that the actual CHD risk patient is correctly predicted as a CHD risk patient, was 837; TN, which means that patients with low CHD risk are correctly predicted as those with low CHD risk, was 218; FP, which means that a person with a low CHD risk is misdiagnosed as a CHD risk patient, was 75; and FN, which means that patients with high CHD risk are incorrectly classified as those with low CHD risk, was 142.

The sensitivity of the optimized-DBN was 0.855, Specificity was 0.744 and accuracy was 0.829. This was confirmed to be superior to the results of the Framingham study, which showed sensitivity as 0.4430, specificity as 0.65, and accuracy as 0.5015. In addition, it performed well when compared with other classifiers. In the case of sensitivity, all of the classifiers are generally high. In particular, the Boltzmann Perceptron Network is slightly higher than suggesting algorithm. But the performance of the proposed algorithm is relatively superior to BPN in other factors such as Specificity, Accuracy and AUC. Even in case of specificity, the performance of the proposed algorithm is superior to other classifiers, but slightly lower than DBN's 0.8209. However, suggesting algorithm is also superior to DBN in other factors. In the case of Accuracy and AUC, it is confirmed that the proposed algorithm shows much better performance than all other classifiers. The performance comparison between the proposed optimized-DBN and other

Table 1: Performance comparison by classifiers

Algorithm	Sensitivity	Accuracy	AUC	
NB	0.8482	0.6385	0.7917	0.736
LR	0.8283	0.6962	0.8003	0.716
BPN	0.8654	0.6211	0.7909	0.701
RF	0.8219	0.6144	0.7720	0.696
DBN	0.7469	0.8209	0.7508	0.570
Framingham	0.4430	0.65	0.5016	0.548
Optimized DBN	0.8550	0.7440	0.8294	0.762

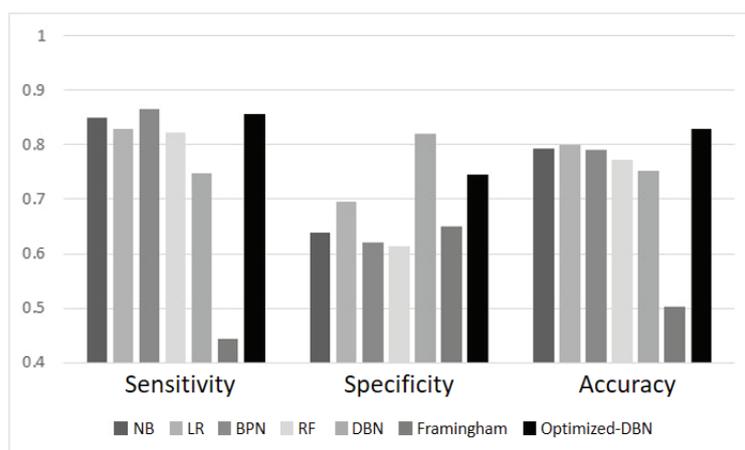


Figure 6: Comparison graph to evaluate the optimized-DBN based CHD risk prediction

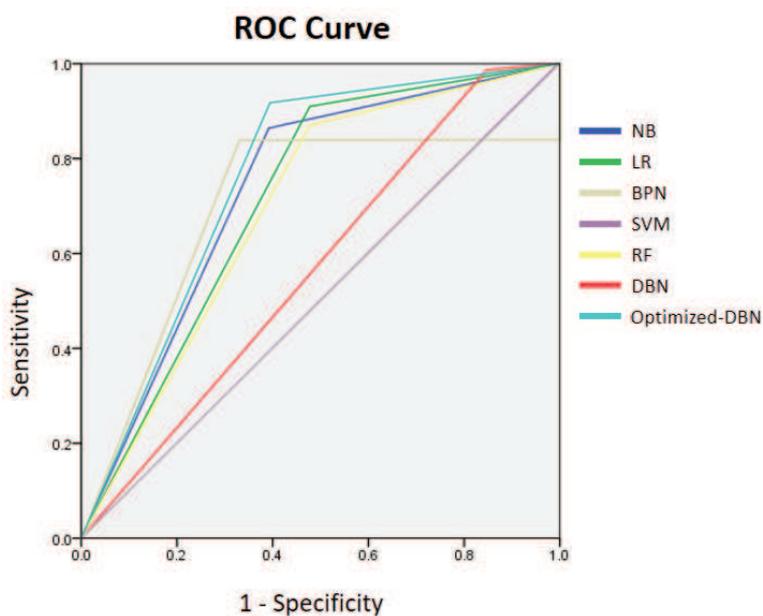


Figure 7: Comparison graph to evaluate the optimized-DBN based CHD risk prediction

prediction algorithms is shown in the figure 6 and figure 7.

The AUC of the optimized-DBN based CHD prediction was 0.762. The Framingham AUC

was 0.546, which confirms that the optimized-DBN based CHD prediction was more accurate. In addition, it performed well when compared with other classifiers such as the support vector machine (SVM) and random forest(RF) which shows 0.501 and 0.696. More precise results are listed the figure 6.

5 Conclusions and future work

The purpose of this study is to develop the Optimized-DBN, a more advanced ANN based prediction model, by combining two ANN algorithms, DBN and GA. And we validate its performance through predicting CHD risk. The results show the improved performance of Optimized-DBN when comparing to not only the conventional DBN but also other major ANN algorithms and Framingham risk score based prediction model which is widely used in healthcare system. This study demonstrates the use of a GA as a tool to draw the optimal parameter values for setting the DBN classifier to improve the performance. It is very useful since it can be applied regardless of datasets because the Optimized-DBN is able to bring the optimal value sets out which are tailored to the data set.

This study will contribute to helping not only medical staffs to make decisions but also patients to prevent CHD by realizing their risk in advance. And it is also able to be applied to any other diseases, if there is enough patient medical data to train and validate the prediction model. However, since this study used only the national health data of Republic of Korea, to more stable and reliable verification, more precise, objective evaluation through further samples is needed. Furthermore, despite its superiority, still, the prediction accuracy is not perfect so we should keep working to develop a more improved way of classification and prediction.

Acknowledgements

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2018-2017-0-01630) supervised by the IITP(Institute for Information & Communications Technology Promotion).

Bibliography

- [1] Blackwell, D. L.; Lucas, J. W. (2014); Summary health statistics for U.S. adults - national health interview survey, *Vital Health statistics*, 10(260), 1–161, 2014.
- [2] Fielding, A. H. (1997); A review of methods for the assessment of prediction errors in conservation presence/absence models *Environmental conservation*, 24(1) 38–49, 1997.
- [3] Freeman J. A.; Skapura, D. M. (1991); *Algorithms, Applications, and Programming Techniques*, Addison-Wesley Publishing Company, 1991.
- [4] Hecht-Nielsen, R. (1992) Theory of the Backpropagation Neural Network *Neural Networks for Perception*, 65–93, 1992.
- [5] Hinton, G. E.; Osindero, S.; Teh, Y.-W. (2006); A fast learning algorithm for deep belief nets *Neural computation*, 18(7), 1527–1554, 2006.
- [6] Eom, J.-H.; Rhee, J.-K. (2006); AptaCDSS-A Cardiovascular Disease Level Prediction and Clinical Decision Support System using Aptamer Biochip, *Korean Institute of Information Scientists and Engineers*, 33, 28–32, 2006.

- [7] Hinton, G.E.; Osindero S.; Teh, Y.W. (2006); A fast learning algorithm for deep belief nets, *Neural computing*, 18(7), 1527-1554, 2006.
- [8] Holland J.H. (1984); Genetic Algorithms and Adaptation, In: Selfridge O.G., Rissland E.L., Arbib M.A. (eds), *Adaptive Control of Ill-Defined Systems. NATO Conference Series (II Systems Science)*, Springer, Boston, MA, 16, 317-333, 1984.
- [9] Ki, S.K.; Lee, S.M. (2014); Voice Activity Detection based on DBN using the Likelihood Ratio *Journal of Rehabilitation Welfare Engineering& Assistive Technology*, 8(3), 145–150, 2014.
- [10] Korea Center for Disease Control and Prevention (2013); *Guidelines for using raw data of Korean National Health and Nutrition Examination Survey - the first survey of the sixth phase (KNHANES VI-1)*, Ministry of Health and Welfare, 2013.
- [11] Korea Encyclopedia Research Center (1996); *Korea Encyclopedia Research Center: Nursery Encyclopedia*, Korea Encyclopedia Research Center, 1996.
- [12] Korea National Health and Nutrition Examination Survey (2013); [Online]. Available: <https://knhanes.cdc.go.kr/knhanes/>
- [13] Lewis, P.O. (1998); A genetic algorithm for maximum-likelihood phylogeny inference using nucleotide sequence data *Molecular Biology and Evolution*, 15(3), 277–283, 1998.
- [14] Mohamed, A.R.; Dahl, G.E.; Hinton, G. (2011); Acoustic modeling using deep belief networks *IEEE Transactions on Audio, Speech, and Language Processing*, 20(1), 14–22, 2011.
- [15] National Institutes of Health (2001); *NIH: National Cholesterol Education Program ATP III Guidelines*, United States National Institutes of Health, 2001.
- [16] Liu, N.; Jiang-ming Kan, J.-M. (2016); Improved Deep Belief Networks and Multi-Feature Fusion for Leaf Identification *Neurocomputing*, 216, 460–467, 2016.
- [17] Park, R.W. (2017); Sharing Clinical Big Data While Protecting Confidentiality and Security: Observational Health Data Sciences and Informatics *Healthcare Informatics Research*, 23(1), 1–3, 2017.
- [18] Townsend, J. T. (1974); Theoretical analysis of an alphabetic confusion matrix *Perception & Psychophysic*, 9(1), 40–50.
- [19] Whitley, D. (1994); A genetic algorithm tutorial *Statistics and computing*, 4(2), 65–85, 1994.
- [20] Wulsin, D.F.; Gupta, J. R.; Mani, R.; Blanco, J. A. (2011); Modeling electroencephalography waveforms with semi-supervised deep belief nets fast classification and anomaly measurement *Journal of neural engineering*, 8(3), 036015, 2011.
- [21] Yan, X.; Chao, T.; Tu, K.; Zhang, Y. (2007); Improving the prediction of human microRNA target genes by using ensemble algorithm *Federation of European Biochemical Societies*, 581(8), 1586–1593, 2007.
- [22] You, H.; Koo, M.-M.; Yi, K.; Nam, K. (2016); The Frequency based Study of the Applicability of DBN Algorithm on Language Acquisition Modeling *The Korean Journal of Cognitive and Biological Psychology*, 28 (4), 617–651, 2016.