# **The methods of duo output neural network ensemble for prediction of coronary heart disease**

### **Wiharto Wiharto, Esti Suryani, and Vicka Cahyawati**  Department of Informatics, Sebelas Maret University, Indonesia



Wiharto Wiharto, Department of Informatics, Sebelas Maret University, Jl. Ir. Sutami, No. 36A, Kentingan, Surakarta, Indonesia. Email: wiharto@staff.uns.ac.id

### **1. INTRODUCTION**

The occurrence of Coronary heart disease (CHD) is hard to predict, therefore the best course of action is an early prediction based on risk factors. Predicted risk of incidence of coronary heart disease in the next 10 years can be done by using Framingham Risk Score or SCORE method. Unfortunately, existing prediction methods always refers to populations for certain countries, such as Framingham risk score which refers to the population in America or the SCORE for population in Europe. This is reinforced in the Kim et al. study [1] which mentions the Framingham Risk Score is not quite suitable for the Korean population. Such discrepancies lead to the development of other appropriate models based on clinical data already available, notably by utilizing the data mining techniques [2].

A number of studies have been conducted, among them by Wiharto et al. [3], proposing predictive models using fuzzy inference system (FIS). Rule and membership function used in FIS refers to the Framingham risk score model. Similar studies were performed by Kim et al. [1] and Kim et al. [4] which suggested a prediction model for the incidence of coronary heart disease by using a fuzzy inference system, with the membership function used referring to the Framingham model risk score as well. Extract rule used in both researches used decision tree algorithm. The extraction rule algorithm used in Kim et al. [1] was C4.5, whereas Kim et al. [4] used the CART. Both studies are equally using clinical data for Korean populations. Other models that use knowledge-based Framingham risk score and PROCAM are performed by Khatibi et al.

ֺ֝

[5]. In the model, the rule and membership function of each model were FRS and PROCAM, respectively, which were subsequently combined by using dempster-shafer algorithm for decision making.

Other modelling work without the Framingham risk score or SCORE as a knowledge base have also been carried out by several researchers. Wiharto et al. [6] used a vector support machine preceded by feature selection using principal component analysis (PCA). The model obtained an accuracy below 80%. Another algorithm that is widely used in classification problem is the artificial neural network (ANN) [7]. Amin et al. [8] combined genetic algorithm to optimize ANN, thereby providing improved performance in predicting coronary heart disease based on risk factors. The use of genetic algorithms can suppress overfitting in the proposed model, even if the amount of data used in training and testing is relatively small. Overfitting can also be suppressed by using data for training in large numbers, as shown by LaFrenier et al. [9]. The study was able to provide a relatively similar performance of accuracy when tested using data training and data testing.

The prediction model of coronary heart disease using artificial neural network was also performed by Yang et al. [10]. The study predicted coronary heart disease in geriatric patients. The model provided better performance than using Framingham risk score. In addition, the performance of the model was also better than using logistic regression, decision tree and linear discriminant analysis (LDA). The artificial neural tissue was also used in the diagnosis of coronary heart disease with a tiered approach [11]. In the study, the prediction at the first level employed risk factor modelled by ANN. The training algorithm used in ANN used Levenberg-Marquard (ANN-LM). The performance obtained using the model was fair, with an accuracy of less than 70%. The fair performance of ANN-LM indicated that overfitting occurred during testing process, thus the performance of training process was better than that of testing. The finding was also supported by [12] which explains that single ANN usage tends to overfit.

Other studies that did not use the knowledge base of the Framingham risk score and SCORE generally added a number of further examinations such as a type of chest pain and troponin. The research by Verma et al. [13] used risk factors coupled with a type of chest pain examination in predicting coronary artery disease. The study tested a number of methods, namely C4.5, Multi-layer Perceptron (MLP) and naive bayesian. The test results showed that C4.5 provides the best performance. Similar studies were carried out by Nithya & Suresh [14], but by using risk factors and troponin examination. The research applied competitive Bat imperialist algorithm for feature selection and support vector machine (SVM) with its kernel-based fuzzy genetic improvement (GF-IKSVM). It also carried out testing using SVM and ISVM, with GF-IKSVM providing the best performance. Prediction using only risk factors was also used in the study of Rajkumar et al. [15] for the prediction of coronary artery disease, but specifically for diabetic patients. The prediction systems used a number of methods, where a neuro-fuzzy classifier was the best method compared to MLP and naive bayesian.

Altan et al. [16] studied the diagnosis of coronary heart disease using the results of an electrocardiogram (ECG) examination. Tthe results of the ECG examination was then classified using deep belief networks. Similar studies combining the risk factors with ECG examination were conducted by a number of researchers namely Alizadehsani et al. [17], Alizadehsani et al. [18], Alkeshuosh et al. [19], and Alizadehsani et al. [20]. Research by Alizadehsani et al. [17] combined symptoms with ECG examination which were classified later using naive bayesian, support vector machine optimization (SMO) and ensemble. Alizadehsani et al. [18] employed the results of examination of risk factors, symptoms and ECG as a feature for diagnosis, while the classification algorithms tested were SMO, SVM, C4.5, Naive bayesian and KNN. CHD diagnosis using the SMO algorithm was found to provide the best performance [17] [18]. Research conducted by Alkeshuosh et al. [19], used particle swarm optimization to build a decision rule in diagnosis from the results of ECG examinations, while the study by Alizadehsani et al. [20] used the Naive Bayesian, SVM and C4.5 algorithm in predicting coronary heart disease.

Many previous studies used ANN to predict CHD; unfortunately, the use of ANN, especially MLP has a disadvantage due to the occurrence of overfitting. The next weakness in terms of techniques used by MLP-ANN for binary classification is in the creation of output thresholds, which must be set before classification. This leads to a lack of clarity in classification since the output of MLP-ANN always becomes uncertain [21]. An alternative to overcome the overfitting in MLP-ANN is with the ensemble system, while the obscurity in the output classification could be overcome with the binary output duo model. The neural network ensemble uses two or more single neural networks that are combined into one with the aggregation model. This study proposes the prediction of coronary heart disease by using Duo Output Ensemble Artificial Neural Network (DOANNE). The model is trained using the Levenberg-Marquard algorithm, therefore the system is called DOANNE-LM. The average aggregation method is used in DOANNE-LM. Accuracy parameter is used for performance evaluation. A performance comparison between single ANN-LM with DOANNE-LM is also conducted by using statistical test, that is by t-test method, with 95% confidence level.

53

### **2. RESEARCH METHOD**

## **2.1. Data**

The data used in this study was obtained from the patient's medical record at RSUD Dr Moewardi, Surakarta, Indonesia. The medical record data has 12 attributes, as shown in Table 1. The data used in this study was divided into two, namely training and testing data. Data for training amounted to 72 patients; 36 patients diagnosed with positive coronary heart disease, and the rest diagnosed with negative coronary heart disease. The testing used 24 test data outside of randomly drawn training data. The data attribute shown in Table 1 has values of varying scales, so the values for each attribute need to be pre-processed, to normalize the data to fit the scale before ANN. The scale refers to the output of the activation function used in ANN, which is in the range 0-1, thus the ANN input was converted to a scale of 0-1. Scaling refers to grouping the factor according to certain criteria, then assigning values from 0-1 to each group in the risk factor. The scaling for all risk factors is shown in Table 1.



#### **2.2. Proposed Method**

The predicted model of coronary heart disease based on risk factors can be modelled as shown in Figure 1 and Figure 2. Figure 1 shows the training model while Figure 2 reveals testing model, both with

DOANNE-LM. DOANNE is an ensemble system with each ANN having two ANN models. The first ANN is trained with specific training data set for specific target of output (e.g. TRUE), while the second ANN is trained with other training data set for complementary target of output (FALSE, in this case). Therefore, the ANN output has two outputs, thus an aggregation method is needed for both outputs. The average was used for aggregation method.



Figure 1. DOANNE-LM training model



Figure 2. DOANNE-LM Testing Model

The DOANNE model has several stages in the training and testing process. The stages are as follows [21]:

- a. Determine the number of neural network pairs to be trained.
- b. For each pair:

1) The first network is trained using true and false target values to predict true and false outputs, i.e.  $T_{\text{target}}^j(x_i)$  and  $F_{\text{target}}^j(x_i)$  . The wrong target value can be obtained by the formula

$$
F_{\text{target}}^j(x_i) = 1 - T_{\text{target}}^j(x_i) \tag{1}
$$

- 2) The second network is trained using wrong and correct target values, i.e.  $F^j_{\text{target}}(x_i)$  and  $T^j_{\text{target}}(x_i)$  to predict the wrong and correct output.
- 3) Finding the average output value is correct with the formula:

$$
T_{avg}^{j}(y_k) = \frac{T_1^{j}(y_k) + (1 - F_1^{j}(y_k))}{2}
$$
 (2)

4) Finding the average value of output wrong with the formula:

$$
F_{avg}^{j}(y_{k}) = \frac{F_{2}^{j}(y_{k}) + (1 - T_{2}^{j}(y_{k}))}{2}
$$
\n(3)

5) Classify results by finding the average value by the formula:

$$
T_{avg}(y_k) = \frac{\sum_{j=1}^{m} T_{avg}^j(y_k)}{m}
$$
\n
$$
\tag{4}
$$

$$
F_{avg}(y_k) = \frac{\sum_{j=1}^{m} F_{avg}^j(y_k)}{m} \tag{5}
$$

Then the value generated by the (4) and (5) is compared based on the input pattern. If  $T_{\text{avg}}(y_k)$  $F_{\text{avg}}(y_k)$  then output is classified as value 1, otherwise it is 0.

The DOANNE model uses a multi-layer perceptron architecture with the Levenberg-Marquard (LM) training algorithm. The LM algorithm is designed using a second derivative approach without the need to compute the Hessian matrix [22]. The LM algorithm has advantages in terms of the speed of the training process, compared to a number of other algorithms [23]. Both predictive models of coronary heart disease with either ANN-LM or DOANNE-LM were measured using accuracy performance parameters. In addition to accuracy performance parameters, statistical tests were also performed to determine the level of significance of system improvements on the overfitting, for ANN-LM as well as DOANNE-LM. The statistical test was done with a 95% confidence level.

### **3. RESULTS AND ANALYSIS**

The prediction system model is proposed using the ensemble system approach. The ensemble system is applied to ANN with duo output (DOANNE-LM). The results of the test prediction model of coronary heart disease using DOANNE-LM, as well as using LM training algorithm (ANN-LM) are shown in Table 2. The test results obtained are from 10 experiments for both predictive models using ANN-LM and DOANNE-LM. The test is done by variation of the number of ensembles, i.e. by multiples of 5, and with a maximum limit of 40 ensembles.

Table 2. Test results for ANN-LM and each number of ensembles

Experiment	Algorithms								
	ANN-LM	Ensemble (Sum of Ensemble) $(\% )$							
	$\frac{1}{2}$		10	15	20	25	30	35	40
	79.167	87.500	87.500	87.500	87.500	91.667	87.500	87.500	87.500
	83.333	87.500	83.333	87.500	83.333	83.333	83.333	87.500	91.667
	87.500	83.333	87.500	91.667	91.667	91.667	91.667	91.667	91.667
4	75.000	83.333	87.500	87.500	83.333	83.333	87.500	83.333	91.667
	79.167	91.667	87.500	87.500	87.500	87.500	87.500	87.500	87.500
6	54.167	83.333	83.333	83.333	87.500	91.667	91.667	91.667	87.500
	70.833	83.333	87.500	83.333	87.500	87.500	87.500	91.667	91.667
8	70.833	87.500	83.333	83.333	83.333	83.333	83.333	83.333	83.333
9	83.333	83.333	87.500	87.500	87.500	83.333	83.333	83.333	83.333
10	87.500	83.333	87.500	87.500	87.500	87.500	87.500	87.500	87.500

*The methods of duo output neural network ensemble for prediction of … (Wiharto Wiharto)* 

The test results of both ANN-LM and DOANNE-LM using training data, for 10 experiments, provide 100% accuracy. The results of the test as shown in Table 2 also express the ability of the DOANNE-LM model to suppress the occurrence of overfitting. The amount of overfitting that occurs for both prediction system model are shown in Table 3. Overfitting on the multi-layer perceptron neural network (MLP-ANN) architecture results in lower MLP-ANN performance. The MLP-ANN model performs well when tested using training data, and low when using data different from training data [24]. The relatively small amount of training data can cause overfitting, while for relatively large amounts of data it can suppress overfitting. This has been observed in a study conducted by LaFreniere [9].

	Training (%)	Testing $(\% )$	Overfitting $(\% )$
ANN-LM	100	77.083	22.917
5 ENSM	100	85.417	14.583
10 ENSM	100	86.250	13.750
15 ENSM	100	86.667	13.333
20 ENSM	100	86.667	13.333
25 ENSM	100	87.083	12.917
30 ENSM	100	87.083	12.917
35 ENSM	100	87.500	12.500
40 ENSM	100	88.333	11.667

Table 3. The performance of the DOANNE-LM model in suppressing the occurrence of overfitting

The test results show the ability of the DOANNE-LM model in suppressing overfitting as shown in Table 3. When using ANN-LM single prediction model, high overfitting occurs, which is 22.917%. Overfitting decreases in the use of DOANNE-LM with an increasing number of ensembles. The decrease in overfitting that occurs when using DOANNE-LM reaches 49.09%. In the ensemble system, one of the most influential parts besides the number of ensembles is the method used for merging. In this study, a combination of ensembles with the average method of the number of ensembles was used. The improvement of DOANNE-LM system performance compared to ANN-LM can be shown from the result of t-test test statistic with 95% confidence level, as shown in Table 4. It shows that p-value <0.05 representing a marked value that there was a significant improvement for the prediction of coronary heart disease using DOANNE-LM, in constrast to ANN-LM.



The proposed prediction system model using DOANNE-LM performs better than previous studies in [6]. Research conducted by Wiharto et al. [6] proposed a prediction system model using a combination of principal component analysis (PCA) with a vector support machine (SVM). The combination of PCA-SVM was only capable of delivering 78.64% accuracy. Subsequent research was conducted by Arabasadi et al. [25] while using a single ANN, and the resulting performance was better than the DOANNE-LM prediction system model. In addition to using ANN as well as combining it with a genetic algorithm (GA-ANN), the performance resulting from the combination is better, with an accuracy of 93.85%. The resulting performance is better due to the research using additional checks of Electrocardiography (ECG), while the DOANNE-LM system model is only based on risk factors. These additional checks lead to increased accuracy. This is highly relevant since the ECG examination is one of the major checks in diagnosing coronary heart disease. It is also reinforced in the another study [11] which used a tiered model. At the second level ECG check was added, either rest or during exercise. The addition of such checks adds to system performance improvements. Comparison with a number of previous studies in full can be shown in Table 5.

Authors	Method	$\beta$ . The second $\beta$ is the second $\alpha$ in the second $\alpha$ Dataset	Sum of data	Accuracy $(\% )$
Kim et al. $[1]$	$FIS+C4.5$	<b>KNHANES</b>	299	69.22
	FIS+CART	<b>KNHANES</b>	748	69.51
	<b>ANN</b>	<b>KNHANES</b>	748	62.78
Kim et al. $[4]$	<b>SVM</b>	<b>KNHANES</b>	748	67.71
	LR	<b>KNHANES</b>	748	63.23
	C5.0	<b>KNHANES</b>	748	53.36
	<b>ANN</b>	<b>KNHANES</b>	335	73.74
Yang et al. $[10]$	Framingham Risk Score	<b>KNHANES</b>	335	64.65
Wiharto et al. [11]	Tiered-ANN	UCI Repository (Risk Factor)	303	61.00
Wiharto et al. [6]	PCM+SVM	Dataset RSUD Dr Moewardi (Risk Factor)	120	78.64
Arabasadi et al. [25]	$GA+ANN$	Risk Factor + ECG	303	93.85
	<b>ANN</b>	Risk Factor +ECG	303	84.62
Abdar et al. [26]	Logistic Regression $+ANN$	UCI Repository (Risk Factor, ECG, Fluoroscopy, Scintigraphy)	303	80.23
Miao et al. [27]	<b>Adaptive Boosting</b> (Ensemble)	UCI Repository (Risk Factor, ECG, Fluoroscopy, Scintigraphy)	303	80.14
Proposed	<b>DOANNE-LM</b>	Dataset RSUD Dr Moewardi (Risk Factor)	96	86.875

Table 5. Comparison of the proposed system model with the previous model

### **4. CONCLUSION**

The prediction model of coronary heart disease using DOANNE-LM has been able to provide significantly better performance than the diagnostic system model with ANN-LM. This is indicated by the results from statistical tests which show the p-value <0.05. The performance of predictive system models with DOANNE-LM is capable of delivering performance with an average of 86.875% accuracy, which is better than many prediction models in published researches. This study has also confirmed that the use of single ANN causes overfitting when used for predicting coronary heart disease (22.917%). The proposed model using DOANNE-LM ensemble system is capable of suppressing overfitting, which has reduced overfitting by 49.09%.

### **REFERENCES**

- [1] Kim JK, Lee JS, Park DK, Lim YS, Lee YH, and Jung EY, "Adaptive mining prediction model for content recommendation to coronary heart disease patients," *Clust. Comput,* vol. 17, no. 3, pp. 881–891, 2014.
- [2] Wiharto W, Herianto H, and Kusnanto H, "System Diagnosis of Coronary Heart Disease Using a Combination of Dimensional Reduction and Data Mining Techniques: A Review," *Indonesian Journal of Electrical Engineering and Computer Science,* vol. 7, no. 2. pp. 514–523, 2017.
- [3] Wiharto W, Herianto H, and Kusnanto H. "Early screening of incidence coronary heart disease based on risk factor using fuzzy rule system," *Int. J. Adv. Appl. Sci,* vol. 4, no. 6, pp. 153–158, 2017.
- [4] Kim J, Lee J, and Lee Y, "Data-Mining-Based Coronary Heart Disease Risk Prediction Model Using Fuzzy Logic and Decision Tree," *Healthc. Inform. Res,* vol. 21, no. 3, pp. 167–174, 2015.
- [5] Khatibi V and Montazer GA, "A fuzzy-evidential hybrid inference engine for coronary heart disease risk assessment," *Expert Syst. Appl,* vol. 37, no. 12, pp. 8536–8542, 2010.
- [6] Wiharto W, Kusnanto H, and Herianto H, "Clinical Decision Support System for Assessment Coronary Heart Disease Based on Risk Factor," *Indian J. Sci. Technol.* vol. 10, no. 22, pp. 1–12, 2017.
- [7] Sug H, "Investigating Better Multi-layer Perceptrons for the Task of Classification," *WSEAS Trans. Comput,* vol. 9, no. 5, pp. 475–480, 2010.
- [8] Amin SU, Agarwal K, and Beg R, "Genetic neural network based data mining in the prediction of heart disease using risk factors," *IEEE Conference on Information & Communication Technologies (ICT). Thuckalay, Tamil Nadu, India,* pp. 1227–1231, 2013.
- [9] LaFreniere D, Zulkernine F, Barber D, and Martin K, "Using machine learning to predict hypertension from a clinical dataset," *IEEE Symposium Series on Computational Intelligence (SSCI),* Athens, Greece, pp. 1–7, 2016.
- [10] Yang J, Lee Y, and Kang UG, "Comparison of Prediction Models for Coronary Heart Diseases in Depression Patients," *Int. J. Multimed. Ubiquitous Eng,* vol. 10, pp. 3, pp. 257–268, 2015.
- [11] Wiharto W, Herianto H, and Kusnanto H, "The Analysis of Performace Model Tiered Artificial Neural Network for Assessment of Coronary Heart Disease," *International Journal Electrical Computer Engineering (IJECE),* vol. 7, no. 4, pp. 2183–2191, 2017.
- [12] Wiharto W, Aziz A, and Permana U, "Improvement of Performance Intrusion Detection System (IDS) Using Artificial Neural Network Ensemble," *J. Theor. Appl. Inf. Technol,* vol. 80, no. 2, pp. 191–201, 2015.
- [13] Verma L, Srivastava S, and Neg, "PC. An intelligent noninvasive model for coronary artery disease detection," *Complex Intell. Syst,* vol. 4, no. 1, pp 11–18, 2018.
- [14] Nithya S and Suresh GD, "Bat imperialist competitive algorithm (BICA) based feature selection and genetic fuzzy based improved kernel support vector machine (GF-IKSVM) classifier for diagnosis of cardiovascular heart disease," *Biomed Res*, vol. 29, pp. S95-S104, 2018.
- [15] Rajkumar R, Ananadakumar K, and Bharathi A, "Risk Level Classification of Coronary Artery Heart Disease in Diabetic Patients using Neuro Fuzzy Classifier," *Int. J. Comput. Intell. Res,* vol. 13, no. 4, pp. 575–582, 2017.
- [16] Altan G, Allahverdi N, and Kutlu Y, "Diagnosis of Coronary Artery Disease Using Deep Belief Networks, "*Eur. J. Eng. Nat. Sci,* vol. 2, no. 1, pp. 29–36, 2017.
- [17] Alizadehsani R *et al,* "Diagnosis of Coronary Artery Disease Using Data Mining Techniques Based on Symptoms and ECG Features," *Eur. J. Sci. Res,* vol. 82, no. 4, pp. 542–553, 2012.
- [18] Alizadehsani R, Hosseini MJ, Sani ZA, Ghandeharioun A, and Boghrati R, "Diagnosis of Coronary Artery Disease Using Cost-Sensitive Algorithms, *IEEE 12th International Conference on Data Mining Workshops*, Brussels, Belgium, pp. 9–16, 2012.
- [19] Alkeshuosh AH, Moghadam MZ, Mansoori IA, and Abdar M, "Using PSO Algorithm for Producing Best Rules in Diagnosis of Heart Disease," I*nternational Conference on Computer and Applications (ICCA)*, Doha, United Arab Emirates, pp. 306–311, 2017.
- [20] Alizadehsani R *et al,* "Non-invasive detection of coronary artery disease in high-risk patients based on the stenosis prediction of separate coronary arteries," *Comput. Methods Programs Biomed,* vol. 162, pp. 119–127, 2018.
- [21] Kraipeerapun P and Amornsamankul S, "Ensemble of Duo Output Neural Networks For Binary Classification," *Int. J. Math. Comput. Simul,* vol. 4, no. 4, pp. 162–170, 2010.
- [22] Wang L, Liu Y, Liu Y, Wang W, Zhao Y, and Yang Z, "Optimization of Hydrogen-fueled Engine Ignition Timing Based on L-M Neural Network Algorithm," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 14, no. 3, pp. 923–932, 2016.
- [23] Mukherjee I and Routroy S, "Comparing the performance of neural networks developed by using Levenberg– Marquardt and Quasi-Newton with the gradient descent algorithm for modelling a multiple response grinding process," *Expert Syst. Appl,* vol. 39, no. 3, pp. 2397–2407, 2012.
- [24] Jong GJ, Huang CS, Yu GJ, and Horng GJ, "Artificial Neural Network Expert System for Integrated Heart Rate Variability," *Wirel. Pers. Commun,* vol. 75, no. 1, pp. 483–509, 2014.
- [25] Arabasadi Z, Alizadehsani R, Roshanzamir M, Moosaei H, and Yarifard AA, "Computer aided decision making for heart disease detection using the hybrid neural network-Genetic algorithm, "*Comput. Methods Programs Biomed,*  vol*.* 141, pp. 19–26, 2017.
- [26] Abdar M, Kalhori SRN, Sutikno T, Subroto IMI, and Arji G, "Comparing Performance of Data Mining Algorithms in Prediction Heart Diseases," I*nternational Journal Electrical Computer Engineering (IJECE)*, vol. 5, no. 6, pp. 1569–1576, 2015.
- [27] Miao KH, Miao JH, and Miao GJ, "Diagnosing Coronary Heart Disease using Ensemble Machine Learning," *Int. J. Adv. Comput. Sci. Appl,* vol. 7, no. 10, pp. 30–39, 2016.