Case Studies on Neural Networks for Prediction in Health/Diseases Problems

Edy Mulyanto*, **Arry Maulana Syarif,**
Fikri Budiman, Khafiizh Hastuti

Department of Computer Science
University Dian Nuswantoro
207 Imam Bonjol Str., Semarang 50131, Indonesia
E-mails: edy.mulyanto@dsn.dinus.ac.id, arry.maulana@dsn.dinus.ac.id, fikri.budiman@dsn.dinus.ac.id, afis @dsn.dinus.ac.id

**Corresponding author*

Received: February 17, 2018 **Accepted: September 12, 2019**

Published: March 31, 2020

Abstract: *Human health is one of topics which require minimum error or even zero tolerance error in all research including in computer science research. Classification in health problems is an interesting and ongoing topic research to solve a problem with minimum error result where neural network has become a popular approach in this area. Types and characteristics of data sources selected to make a prediction for seven health disease problems were described in this paper. The summary of the report is expected to be stimulation for next researches interest in selecting a method or a technique which is appropriate with the health problem to solve.*

Keywords: *Neural network, Classification, Prediction, Health.*

Introduction

Human health is one of computer science research topics which requires the minimum error or even zero tolerance error. The development of artificial intelligence and machine learning has stimulated many researchers to build various systems which can support many health services with minimum error in result. A classifier system to predict certain disease such as classifier for heart disease risk prediction, breast cancer risk prediction, and others, has become an interesting research to conduct. Neural network is an approach often used to build a classifier system including in solving health problems.

Neural network firstly introduced by Warren McCulloch and Walter Pits in 1943 is a model of the information process capabilities using mathematical analysis inspired by biological learning system of nervous system [5, 7, 12, 15]. Neural network provides nonlinear algorithms for feature extraction using hidden layer and classification using multiperceptron [2]. There are several types of neural network models, such as: single layer perceptron model which is the simplest form of neural network, multilayer perceptron model which has input layer, hidden layer and output layer as found in Feedforward Neural Network (FNN), Convolutional Neural Networks (CNNs) designed to mimic the structure of the animal visual cortex, and most frequently used for image processing and computer vision, Recurrent Neural Networks (RNNs) where the connections between units form a directed cycle [3].

Neural network is one of approaches effectively proven for classification in various problems. This research aims to summarize the use of neural network in health problems based on some published researches. Classification for health problems is an interesting topic in research with many challenges and innovations where neural network has been a popular approach in this area. The goal of this research is to synthesize various methods and techniques of data pre-processing and classification procedures used in some published researches starting from years 2012 randomly selected. The analysis of the selected research publications was conducted in research methodology categorized based on similarity in method or subject where each category was represented with at least two literatures. The study conducted in this paper does not represent all health problems, types of neural networks and methodology used in neural network research, but the summary of this study is expected to be stimulation for further research.

Classification using patient record data

Many classifiers are built based on patient record data used as features. A model for predicting Surgery-Related Pressure Injury (SRPI) in cardiovascular surgical patients using neural network was proposed by [6]. SRPI occurs due to a shearing injury to a patient after surgery, and the injury depends on the individual's intrinsic ability in tolerating it. The demographic and pertinent clinical data, including age, gender, weight disease category, smoking status, and diabetes mellitus are used as the inputs. There was an exclusion criterion for patients with specific characteristic. The stage of pressure injury was determined based on classification of National Pressure Ulcer Advisory Panel-European Pressure Ulcer Advisory Panel-Pan Pacific Pressure Injury Alliance (NPUAP-EPUAP-PPPIA). The possible risk factors were analyzed using statistic tests of univariate student's *t* test and X^2 test, and then the results were entered into a feedforward back propagation neural network. The training data used 70% of cases, and testing data used 30% cases. The potential risk factors identified by univariate analysis were used for the input layer. The network performance was evaluated using Receiver Operating Characteristic (ROC). The distribution of *C*-index which is the value of area under the receiver operating characteristic curve (AUC) that represents the prediction ability is divided into three classes [17]: value ranges 0.5 to 0.7 have mild accuracy, 0.7-0 has moderate accuracy, and 0.9-1 has high accuracy. The experiment results were reported that the performance of prediction ability of the neural network model was in moderate class a *C*-index of 0.815.

A classifier using neural network was developed by [22] to predict a heart disease risk. Cleveland Clinical Foundation Heart Disease dataset with 13 attributes that is significantly relevant to cause a heart disease was used as inputs to the network. The dataset contained 297 samples divided to 138 samples with possibility of developing a heart disease, and 159 samples with no risk of developing heart disease. Min-Max normalization technique was used to transform data. Non-numeric and continuous values data were transformed into numeric data, and specific ranges value. After several experiments to define training algorithm and networks architecture, a feedforward neural network with 13 units in the input layer, seven units in the hidden layer and one unit in the output layer, and back propagation learning algorithm was used for the neural network model. Levenberg-Marquardt back propagation supervised training algorithm was chosen for training, validation and testing with the ratio of data 80:10:10. Performance evaluation was claimed achieving 100% for sensitivity, specificity, precision and accuracy.

Data mining classification technique and Ensemble of Online Sequential Extreme Learning Machine (EOS-ELM) was proposed by [9] to develop a decision support system for prediction of patients with heart disease. The database was collected from Cleveland Heart Disease Dataset (CHDD), and 909 records with 15 attributes including age, sex, chest pain type, etc, were used as the inputs. The training data used 455 records, and testing data used 454 records randomly selected. A data preprocessing phase was conducted by removing missing fields and outliers, and normalizing data. The Modified Genetic Algorithm (MGA) was used to reduce the number of features. The Online Sequential Extreme Learning Machine (OS-ELM) is used for a batch learning, which its hidden units parameters are randomly selected and the output weights are analytically determined. The EOS-ELM which consists of many OS-ELM networks can work better than individual OS-ELM network. The parameters randomly generated make distinction for each OS-ELM network in the ensemble, so some of networks may adapt faster and better than others. The target to be predicted was a diagnosis with multi-classes of absence, moderate and presence of heart disease. Compared to the other classification methods for the accuracy of the prediction, and evaluation using partition entropy coefficient and *V*-measure, the result was reported that EOS-ELM outperformed naive bayes, decision tree, and artificial neural network, and the MGA used in a feedforward neural network to optimize dimensionality reduction in the process of dataset attributes feature extraction significantly speeds up the prediction task with promising accuracy performance achievement.

A feedforward neural network was reported achieving accuracy at moderate class in predicting SRPI based on *C*-index, a measurement proposed by [17]. Another one [22] was claimed achieving accuracy of 100% in predicting a heart disease risk, and also performance of sensitivity, specificity, and precision did, where the achievement can be classified as high accuracy in *C*-index measurement. Considering a perfect achievement of performance evaluation by [22] with the number of dataset, an addition of performance evaluations should be conducted, such as using cross validation or *K*-fold cross validation, in which data for training and testing are rotated. This evaluation can make sure whether the performance evaluation still achieves at 100% in every number of *K* of data rotation. Although the two networks were implemented in different cases, both of them used patient record data pre-processed before the data were sent to the network, in which the attributes were selected based on statistical analysis by [6] and based on information of other researchers by [6, 22].

Data mining technique was combined with a neural network classifier by adding MGA in a feedforward neural network [9] to reduce the dimension of dataset attributes. In other researches, data mining technique, such as association rules, can be used to select the attributes which are potential and relevant. Apriori algorithm has been used in various classification researches to select the best attributes based on their frequent appearance, for examples, in works by [20, 21], apriori algorithm was used to select valuable attributes for Naïve Bayes classifier.

Classification using a large patient record data

A classifier using a large patient record data which contains more than 15 000 records was proposed by [8, 18]. A large record data often deals with imbalanced dataset, such a research conducted by [8] that develop a neural network classifier to predict complications following posterior lumbar fusion using a dataset containing information of 22 629 patients collected from The American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP) Database. The network used 70% of the dataset for training the model, in which the data were distributed for training, validation and testing in a 2:1:1 ratio, and 30% for hold out data as post-training evaluation of the model. The low incidence of complications was overcome using Adaptive Synthetic (ADASYN) sampling approach for learning from imbalanced dataset by generating positive complications as a means. There were 13 input features such as sex, age, ethnicity, diabetes, etc. Six features with the greatest regression coefficient magnitudes were used as input variables for mortality, Venous Thromboembolism (VTE), cardiac complications, and wound complications. The multiple networks were used for a large class imbalance problem. The final testing of the performance was conducted using area under the receiver operating characteristic curve, and the neural network classifier was compared with Logistic Regression (LR) using hold-out data. The report of the experiment results was that neural network model was more sensitive than the LR model in predicting the mortality and wound complications, while LR had greater specificity for mortality and wound complications. A prediction for large class-imbalance inherent with hold-out data for evaluation using neural network was better than LR.

A CNN classifier was developed by [18] to predict "onset" or "weaning" of multiple invasive interventions for supporting the needs of Intensive Care Units (ICUs) needs where clinicians has to make a real-time decision of clinical interventions based on the data of the patient that are complicated. The data collected from the Multiparameter Intelligent Monitoring in Intensive Care (MIMIC-III v1.4) database contained 34 148 patients, in which each record was filtered based on their age and ICU stay duration. There were five target interventions: invasive ventilation, Non-Invasive (NI) ventilation, vasopressors, colloid boluses, and crystalloid boluses, and four possible outputs to predict: onset, wean, stay on, or stay off. The ratio of the train, validation, and test data is 70:10:20. The weighted loss function during optimization was used for class imbalances. The training was stopped based on AUC performance on the validation set. The experiment was conducted by comparing CNN model with Long Short-Term Memory (LSTM) model, and the result was reported that both CNN and LSTM model achieved state-of-the art prediction result for the intervention tasks.

FNN, LSTM, and CNN used in two researches described above were reported achieving good results as classifiers for large data, where a large record data often deals with imbalance problem. The low incidence of complications was overcome using ADASYN for learning from imbalanced dataset, but imbalance problem still happened after ADASYN [8]. On the other hand, class imbalances were overcome with weighted loss function [18]. Instead using multiple networks, combining the method used by [8] with weight loss function used by [18] to overcome the imbalance problem might get a better solution for imbalance problem.

Classification using thermal imaging

Neural network and thermal imaging technique were used for diagnostic of diabetes [1, 14]. A sequence of thermal images of temperature distributions across the sole area of the patient foot was collected for dataset [1]. Image was recorded using a thermal imaging camera with certain specification. Four temperature spots are used as inputs for the input layer, and the output layer has a unit for the diagnosis. Dataset which contains realistic data from 90 patients was divided into approximately 60% for training data, and 40% for testing data. Network was evaluated using Root Mean Square Error (RMSE). The problem of defining number of units in the hidden layer often occurs; network performance improvement can be obtained by increasing number of units in the hidden layer, but this may increase over-fitting. The network evaluation result was used to define an appropriate number of units in the hidden layer where the model with 10 units was the lowest error for testing data. Neural network was compared with other methods to for diabetes diagnostic prediction [14]. Twenty five thermal images across the patient foot were clustered using *k*-means, and then extracted using Gray Level Co-occurrence Matrix (GLCM), and finally classified by comparing three methods: Probabilistic Neural Network (PNN), *K*-nearest Neighbor Network (KNN), and Support Vector Machine (SVM). The reported result showed that the accuracy of KNN and SVM outperformed PNN where PNN achieved accuracy of 61%.

Both [1] and [14] used foot thermal image with different foot area regions for diagnostic of diabetes. Comparing to the area to be captured, [1] used sole area of the patient foot which was smaller than the area captured by [14]. This makes computational cost proposed by [1] was more efficiently than [14] did. The performance evaluation conducted by [1] tends to find an appropriate unit number in hidden layer to get the lowest error without explaining the prediction accuracy, while [14] did not clearly explain the method and architecture of the network. The different data source or a diagnostic of diabetes used by these two researches lead to the question of “which gives the better accuracy”; are the thermal images data from across the sole area can achieve a better prediction than data from across the foot area or vice versa? A prediction for health disease problem should have zero tolerance for error prediction so the use of all the two areas as dataset might be done if it can increase the prediction accuracy despite consequently increasing the computational cost. Therefore it is interesting to more investigate the use of thermal image for diagnostic of diabetes by exploring and synthesizing ideas proposed by [1] and [14].

The use of thermal imaging technique in breast cancer detection using neural network was conducted by [10, 16, 23]. A dataset containing 94 breast images with 48 images of normal breasts and 46 images with some anomaly collected from 47 patients was used by [10] to develop a classifier for breast cancer detection. The dataset was collected from Database for Mastology Research (DMR). The Canny edge detector was used to create mask in order to segment the region of interest into right breast and left breast. Asymmetry analysis was used to identify a normal or abnormal breast. The pixel intensity distribution in right breast and left breast was analyzed using histogram with X-axis represented radiant heat, and Y-axis represented number of pixels for each intensity value. The analysis showed comparison of histogram data of right breast and left breast. The statistical measurements for diagnosis of abnormalities in the breast consisted of eight features: variance, standard deviation, skewness, kurtosis, entropy, range and median. The classifier model used neural network with non-linear classification and back propagation algorithm. The network consisted of eight units in the input layer, five units in the hidden layer, and two units in the output layer. Network was trained using 50 images containing 25 images with some anomaly and 25 normal images. The network performance was evaluated using confusion matrix, and the reported result showed that the network achieved 87% in sensitivity, 83% in specificity and 85% in accuracy.

Patient movement during examination by dynamic thermography was concerned by [16] in order to select the best thermal image for a breast cancer prediction. A process called registration images were performed by considering two images of the same scene to select the best overlap between them. Further, a thermal signal from dynamic sequence of 20 images of a particular patient with X-axis represented time series, Y-axis represented thermal signal index of each mean temperature square of the image, and Z-axis represented temperature was achieved by segmenting breast region of the first image of the sequence, and dividing the region into a grid of 3×3 pixels squares. Signal mobility and signal complexity was used as features. *K*-means algorithm was used to cluster images into healthy patients with similarity in their thermal signal, and sick patients. The experiment was conducted using images of 22 patients which consisted of 11 sick images and 11 healthy images. Three classification methods were used in the experiment. The results were reported that neural network classifier achieved 90.91% in accuracy, Bayesian network classifier achieved 86.36%, and decision tree classifier achieved 90.91%.

Neural network combined with Genetic algorithm was used for breast cancer diagnosis prediction by [23]. Dataset contained infrared images taken from 200 women with ages 18-35 where 15 subjects had abnormal image of their breast. Eight features were used to classify: age, mean, differences between right breast and left breast, variance, skewness, kurtosis, entropy and thermal pattern. Diagnostic parameters of the features was used in the network to increase the convergence by normalize the input data before entering the network. The model used a feedforward neural network with back propagation learning algorithm containing eight units in the input layer, six units in the hidden layer, and one output. Genetic algorithm was used to choose optimal values of parameters used for features. The result was reported that the model achieved the best three parameters out of eight diagnostic parameters, which are kurtosis, differences between right breast and left breast, and thermal pattern with 50% sensitivity, 75% specificity and 70% accuracy.

Researchers [10, 16, 23] used thermal imaging technique in breast cancer detection. Static thermography was used by [10], while dynamic thermography was used by [16]. Neural network was combined with Genetic algorithm to optimize a breast cancer risk prediction process by reducing the features dimension [23]. The genetic algorithm used in a feedforward neural network to optimize dimensionality reduction in the feature extraction process can select the best three diagnostic factors from eight diagnostic parameters of thermal image data with promising accuracy performance achievement. Conditions and temperature of an examining room affects the thermal contrast of captured images in the static thermography, while dynamic thermography does not depend on them by selecting the best image from an image sequence of a patient [10]. Based on the accuracy performance in these two researches, the dynamic thermal imaging outperformed the static thermal imaging. However, it is interesting to compare the static and dynamic thermal imaging using same data.

Classification using electroencephalogram data

A feedforward neural network was used to build a classifier with inputs collected from Electro Encepalogram (EEG) data which contain electrical activity of the brain record used by Brain-Computer Interface (BCI) to communicate with human brain [4]. The task of classifier was to differentiate mental classes. Discrete Cosine Transform (DCT) was used as feature extraction method to solve problems of dataset size, high frequency noise, training time for the network. The network used 10 units in the input layer which represented first 10 DCT coefficients, and two units in the output layer. The number of units in the hidden layer was set after conducting some trials, and 200 units was the best number based on minimum error. The dataset was collected from dataset IV from BCI competition II datasets containing taken from a healthy subject during no-feedback session, and divided into training, validation and testing data with ratio of 70%, 15% and 15%. The performance evaluation conducted using confusion matrix reported that the prediction accuracy achieved 85.1%.

A CNN with initial learning rate of 0.01, momentum of 0.1, batch size of 128, 200 epoch and architecture I(22×22) C(4×4.16) RPFSO was used to build a classifier using EEG data as inputs with four classes to predict [19]. The dataset was collected from BCI signal Dataset 2a consisted of sample rate recorded EEG signal for 9 healthy test subjects. After comparing some feature extraction methods, Fast Fourier Transform Energy Map (FFTEM) worked by generating a 2D feature map from EEG, and transforming EEG channel signal into frequency domain to form a single row in the feature map was reported as the best feature extraction method with 70% in training and 68% in testing.

A feedforward neural network classifier using EEG data as inputs was used to predict two classes targets, and a convolutional neural network classifier was built to predict four classes targets. The feature extraction method selection was an important part in both experiments, where DCT was used to compress EEG signal data before they were used as inputs for the FNN classifier [4], and FFTEM reported as the best method outperformed some feature extraction methods including DCT in generating 2D feature map for the CNN classifier [19].

The EEG signal data pre-processing including feature extraction method selection is an interesting topic for further research. Extracted features which are more suitable for the network needs in solving the classification task can increase the accuracy performance, and decrease the time for training.

Classification using OCT scans data

A feedforward neural network with back propagation learning algorithm was used by [11] to predict a diagnostic precision of the spectralis Optical Coherence Tomography (OCT) for Retinal Nerve Fiber Layer (RNFL) measurement evaluation used to predict Multiple Sclerosis (MS) divided into two sub-groups: eyes with antecedent Optic Neuritis (ON) and without previous ON attack, or healthy. Dataset after excluding samples which did not complete a required test contained 105 consecutive healthy individuals and 112 patients with MS. OCT tests were performed using spectralis device where the signal strength were indicated using a blue quality bar in the image. A minimum quality score was set to select images. Input layer used relationships of age, refraction, and OCT thickness measurements as inputs, and output layer contained target of healthy, MS with ON, and MS with non-ON. The number of units in hidden layer was set by several trials. The use of collected data was maximized using 10-fold cross validation method. Chi square test was performed to evaluate the significant differences in the sex distribution between MS and healthy. Mann-Whitney test was performed to evaluate the significant differences in the distribution of age, visual acuity, and refractive error. The experiment results in predicting MS and healthy was reported that the network achieved performance of sensitivity, specificity, and precision of 89.3%, 87.6%, and 88.5%, and sensitivity and specificity in predicting MS with ON and MS with non-ON were 84.5% and 83.2%.

A deep CNN was used by [13] to predict treatment indication based on central retinal Optical coherence tomography (OCT) scans. Intravitreal injections with anti-vascular endothelial growth factor (anti-VEGF) medications are the standard of care for their indications. The OCT scans of the central retina are used to determine anti-vascular endothelial growth factor (anti-VEGF) indications. The images of OCT scans were exported to obtain the star pattern in the DICOM file format. The reverse engineering technique was used to export large quantities of images. Further, the image series that did not appropriate with the star pattern were excluded from the database. The metadata of the image containing a unique institutional patient identification number was extracted and stored in a database. The data were then classified based on the acquisition date of the OCT images into two types of datasets which are injection and no injection. The ratio of 9:1 was used for training and test datasets. The CNN was modeled using the Caffe framework (a deep learning framework developed by Berkeley Artificial Intelligence Research, University of California) and the architecture was developed using GoogLeNet. Batch normalization was added to increase the learning ability by normalizing input activation for each layer across the input batch of images. A supervised learning approach with the back propagation learning algorithm was used to train the networks. Softmax function which turns numbers into probabilities that sum to one was used as probabilities of the input images belonging to the two groups. The dataset contained 30 567 image series representing 183 402 individual OCT scans. Based on the evaluation of the training, the model had the prediction accuracy of 94.5% for correctly labeling images in the test dataset. The prediction accuracy was further conducting by measuring the prediction accuracy for the validation dataset created by removing scans from patients who contributed images to the training dataset, and the result showed that the model had the prediction accuracy of 95.5%. The networks performance was evaluated using ROC curve, and the result was reported that the AUC was

0.968 for the validation dataset. A second ROC was generated by adding Softmax outputs for the six images per image series, and the AUC was 0.988 for the validation dataset. The calculation of optimal cutoffs for the per-image series classification in the validation dataset had a specificity of 94.1% and a sensitivity of 94.2%.

OCT scans data was used in a feedforward neural network classifier to predict MS and healthy, and sub-class of MS with ON and non-ON, and in a CNN classifier to predict treatment indication with two targets of injection and no injection. The striking difference between them is the number of dataset. Despite the different cases to predict between them, given an accuracy value which is set of 85% as high accuracy, then both of them achieved high accuracy for performance network except for the sub-class of MS prediction. It is interesting to investigate the dataset quantity which necessarily is required to obtain high accuracy in prediction using OCT scans data, or in the topic of health problems prediction.

Conclusion

This paper studies on seven health problems tried to predict using neural network approach with different methods and techniques, which are predictions of surgery-related pressure injury, heart disease, complications following posterior lumbar fusion, multiple invasive interventions, diabetes, breast cancer, mental classes, multiple sclerosis and treatment indication, as described in Table 1.

Table 1. Description of neural network approach for predicting nine health problems

Health problems	**Datasets**		**Architectures**	**Ref.**
SRPI risk	Patient record data	149 samples Training 70% Testing 30%	FNN (4-8-1)	[6]
Heart disease		297 samples Training 80% Validation 10% Test 10%	FNN (13-7-1)	[22]
		909 samples Training 50% Test 50%	FNN (architecture was not described) + MGA	[9]
Complications following posterior lumbar fusion	Large patient record data	22 629 samples Training 70% (with ratio at 2:1:1 for training, validation and testing). Hold out 30%	FNN (3 networks with 6-4-1each)	[8]
Multiple invasive interventions risk		34 148 samples Training 70% Validation 10% Test 20%	LSTM (2 hidden layers with 512 nodes each) CNN (3 different granularities, max-pools and 2 fully connected layers)	[18]

Diabetes	Thermal imaging data	90 samples Training 60% Test 40%	FNN (10-4-1)	[1]
		25 samples Training, validation or test data were not described	FNN (architecture was not described)	[14]
Breast cancer		94 samples Training 53% Test 100%	FNN (8-5-2)	[10]
		22 samples Training, validation or test data were not described	FNN (architecture was not described)	[16]
		200 samples. Training, validation or test data were not described	FNN (8-6-1) + GA	[23]
Mental classes	EEG data	416 samples (1000 Hz) Training 75% Validation 15% Test 15%	FNN (10-200-1)	[4]
		22 samples (250 Hz) Training 70% Test 68%	CNN (I(22x22) C(4×4.16) RPFSO)	[19]
Multiple sclerosis	OCT scan data	217 samples	FNN (architecture was not described)	[11]
Treatment indication with anti-VEGF medications.		171 024samples Training 90% Test 10%	CNN (GoogLeNet inception architecture)	[13]

There are cases which use different sample data quantity where a dataset containing more than 15 000 samples is categorized as a large dataset [8]. A large dataset FNN model was used in the cases of SRPI prediction using by [8], multiple invasive interventions prediction by [18] and treatment indication by [13]. In contrast with patient record data which are used by [6] and [22], and some others, a large patient record data was used for inputs by [8] and [18]. This condition is also found in [13] that used a large OCT scans data for inputs while [11] used a small OCT scans data, and an image processing technique using thermal images by [10] did.

There are some basic questions that come regarding to the relation between network performance and sample data quantity, does a large dataset significantly increase the network performance in solving health problems prediction? Are there some specific health problems cases which need a large dataset? Data quantity used as inputs affects the network model selection, data pre-processing method, network performance and also method combination to optimize the network performance such as works by [9] and [23]. Therefore the number of samples data used for the network should be emphasized with objective arguments. These some basic questions with a large variety of health problems including those that are not mentioned in this paper leads the needs of neural network classification mapping for health problems.

The result can be a useful guidance for researchers who try to solve health problems using neural network classification.

References

1. Abdulshahed A. M., F. M. Alabyad, H. A. Goohe, M. A. Saed (2017). Early Detection of Diabetes Using Thermography and Artificial Neural Networks, International Journal of Computational and Neural Engineering, 4(2), 71-75.
2. Basu J. K., D. Bhattacharyya, T. Kim (2010). Use of Artificial Neural Network in Pattern Recognition, International Journal of Software Engineering and Its Applications, 4(2), 23-33.
3. Beysolow II T. (2017). Introduction to Deep Learning Using R, Apress, California, USA.
4. Birvinskas D., V. Jusas, I. Martišius, R. Damaševičius (2013). Data Compression of EEG Signals for Artificial Neural Network Classification, Information Technology and Control, 42(3), 238-241.
5. Brown A. (1991). Nerve Cells and Nervous Systems, Springer-Verlag, Berlin.
6. Chen H. L., S. J. Yu, Y. X. Xu, S. Q. Yu, J. Q. Zhang, J. Y. Zhao, P. Liu, B. Zhu (2017). Artificial Neural Network: A Method for Prediction of Surgery-related Pressure Injury in Cardiovascular Surgical Patients, Journal of Wound, Ostomy and Continence Nursing, 1-5.
7. Gurney K. (1997). An Introduction to Neural Networks, UCL Press, London.
8. Kim J. S., R. K. Merrill, V. A. Arvind, D. Kaji, S. D. Pasik, C. C. Nwachukwu, L. Vargas, N. S. Osman, E. K. Oermann, J. M. Caridi, S. K. Cho (2018). Examining the Ability of Artificial Neural Networks Machine Learning Models to Accurately Predict Complications Following Posterior Lumbar Spine Fusion, Spine (Phila Pa 1976), 43(12), 853-860.
9. Kumar A. V. S. (2017). Ensemble Online Sequential Extreme Learning Machine and Swarm Intelligent Based Feature Selection for Cleveland Heart Disease Prediction System, International Journal of Advanced Trends in Computer Science and Engineering, 6(5), 84-91.
10. Lessa V., M. Marengoni (2016). Applying Artificial Neural Network for the Classification of Breast Cancer Using Infrared Thermographic Images, Lecture Notes in Computer Science, 9972, 429-438.
11. Martin E. G., R. Herrero, M. P. Bambo, J. R. Ara, J. Martin, V. Polo, J. M. Larrosa, J. G. Feijoo, L. E. Pablo (2015). Artificial Neural Network Techniques to Improve the Ability of Optical Coherence Tomography to Detect Optic Neuritis, Seminars in Ophthalmology, 30(1), 11-19.
12. Mitchell T. M. (1997). Machine Learning, McGraw-Hill Science, New York.
13. Prahs P., V. Radeck, C. Mayer, Y. Cvetkov, N. Cvetkova, H. Helbig, D. Märker (2018). OCT-based Deep Learning Algorithm for the Evaluation of Treatment Indication with Anti-vascular Endothelial Growth Factor Medications, Graefes Arch Clin Exp Ophthalmol, 256(1), 91-98.
14. Purnima S., P. S. Angelin, R. Priyanka, G. Subasri, R. Venkatesh (2017). Automated Detection of Diabetic Foot Using Thermal Images by Neural Network Classifiers, International Journal of Emerging Trends in Science and Technology, 04(05), 5183-5188.
15. Rojas R. (1996). Neural Networks: A Systematic Introduction, Springer, New York.
16. Silva L. F., G. O. Sequeiros, M. L. O. Santos, C. A. P. Fontes, D. C. M. Saade, A. Conci (2015). Thermal Signal Analysis for Breast Cancer Risk Verification, In: Medinfo 2015: eHealth-enabled Health, Sarkar I.N. et al. (Eds.), IMIA and IOS Press, 746-750.
17. Streiner D. L., J. Cairney (2007). What's under the ROC? An Introduction to Receiver Operating Characteristics Curves, The Canadian Journal of Psychiatry, 52, 121-128.

18. Suresh H., N. Hunt, A. Johnson, L. A. Celi, P. Szolovits, M. Ghassemi (2017). Clinical Intervention Prediction and Understanding with Deep Neural Networks, Proceedings of Machine Learning in Healthcare Conference, 68, 322-337.
19. Uktveris T., V. Jusas (2017). Application of Convolutional Neural Networks to Four-class Motor Imagery Classification Problem, Information Technology and Control, 46(2), 260-273.
20. Xiao S., Y. Hu, J. Han, R. Zhou, J. Wen (2016). Bayesian Networks-based Association Rules and Knowledge Reuse in Maintenance Decision-making of Industrial Product-Service Systems, Procedia CIRP, 47, 198-203.
21. Yang T., K. Qian, D.-C. T. Lo, Y. Xie, Y. Shi, L. Tao (2016). Improve the Prediction Accuracy of Naïve Bayes Classifier with Association Rule Mining, 2016 IEEE 2nd International Conference on Big Data Security on Cloud, IEEE International Conference on High Performance and Smart Computing, IEEE International Conference on Intelligent Data and Security, 129-133, doi: 10.1109/BigDataSecurity-HPSC-IDS.2016.38.
22. Yazdani A., K. Ramakrishnan (2017). Performance Evaluation of Artificial Neural Network Models for the Prediction of the Risk of Heart Disease, IEEE International Conference on Systems, Man, and Cybernetics Banff Center, Banff, Canada, 2778-2783.
23. Zadeh H. G., J. Haddadnia, M. Hashemian, K. Hassanpour (2012). Diagnosis of Breast Cancer Using a Combination of Genetic Algorithm and Artificial Neural Network in Medical Infrared Thermal Imaging, Iranian Journal of Medical Physics, 9(4), 265-274.

Edy Mulyanto, S.Si., M.Kom.
E-mail: edy.mulyanto@dsn.dinus.ac.id

Edy Mulyanto received his M.Kom. Degree from Informatics Engineering Departement of Sekolah Tinggi Teknik Informatik Benarif Indonesia. He is currently teaching at Faculty of Computer Science of University Dian Nuswantoro. His research interests are data mining, neural network and artificial intelligent.

Arry M. Syarif, S.Si., M.Kom.
E-mail: arry.maulana@dsn.dinus.ac.id

Arry Maulana Syarif received his M.Kom. Degree from Informatics Engineering Department of Dian Nuswantoro University Semarang. He is currently teaching at Faculty of Computer Science of University Dian Nuswantoro, and studying at doctoral program of Computer Sciences and Electronics Department of University Gadjah Mada. His research interests are data mining, neural network and artificial intelligent.

Fikri Budiman, Ph.D.
E-mail: fikri.budiman@dsn.dinus.ac.id

Fikri Budiman received his M.Kom. and Ph.D. Degrees from Informatics Engineering Department of University Gunadarma. His research interests are image processing and artificial intelligent.

Khafiizh Hastuti, Ph.D.
E-mail: afis @dsn.dinus.ac.id

Khafiizh Hastuti received her M.Kom. degree from Informatics Engineering Department of University Dian Nuswantoro, and Ph.D. Degree from Computer Sciences and Electronics Department of University Gadjah Mada. Her research interests are artificial intelligent, data mining, and software engineering.