2018 Yasin_2018_J._Phys.%3A_Conf._Ser._1025_012095

by Hasbi Yasin

Classification of Company Performance using Weighted Probabilistic Neural Network

Classification of Company Performance using Weighted Probabilistic Neural Network

Hasbi Yasin[1], Adi Waridi Basyiruddin Arifin[1], Budi Warsito[1]

[1]Department of Statistics, Faculty of Science and Mathematics, Diponegoro University
Jl. Prof. Soedharto, SH, Tembalang, Semarang 50275, Indonesia
E-mail: hasbiyasin@live.undip.ac.id

Abstract. Classification of company performance can be judged by looking at its financial status, whether good or bad state. Classification of company performance can be achieved by some approach, either parametric or non-parametric. Neural Network is one of non-parametric methods. One of Artificial Neural Network (ANN) models is Probabilistic Neural Network (PNN). PNN consists of four layers, i.e. input layer, pattern layer, addition layer, and output layer. The distance function used is the euclidean distance and each class share the same values as their weights. In this study used PNN that has been modified on the weighting process between the pattern layer and the addition layer by involving the calculation of the mahalanobis distance. This model is called the Weighted Probabilistic Neural Network (WPNN). The results show that the company's performance modeling with the WPNN model has a very high accuracy that reaches 100%.

1. Introduction

Classification of company performance can be judged by looking at its financial status, whether poor or good state. A company can be said has a good state in term of financial when it can survive in any kind of economic conditions, by looking at its ability to fulfill its financial obligations, and carry out its operations constantly and maintain its continuity to expand its company from time being. In order to achieve that objective, its management or executive company are bound to confront many problems in technical section, administrative, or financial [1].

A financial report can be understood by many parties, when it has been analyzed and interpreted. Analyzing a financial report is an activity that designated to gather and interpret any kind of information from those reports for certain purpose. The output from that analyze will be used by its management for evaluate its performance and achievement then making policy in order to be a better company in the future. While for investors, they can use it for investment strategy and planning. For creditor purpose, it will affect their lending policy in the next period. In order to classified company performance whether it has a good finance or poor finance, can be measured and calculated by using Decree of the Ministry of Finance of the Republic of Indonesia No 740/KMK.00/1989 from its measure performance score attachment that has been revised in Decree of the Ministry of Finance of the Republic of Indonesia No 826/KMK.013/1992 where liquidity ratio, profitability ratio, and solvency ratio as its fundamental measurement.

Classification of company performance can be conducted with some approach, either parametric or non-parametric. ANN is one of the non-parametric methods for classification. An artificial neural network is an information-processing system that has certain performance characteristics in common

with biological neural networks [2]. Probabilistic Neural Network (PNN) is an ANN model for classifying purpose. PNN consists of four layers, i.e. input layer, pattern layer, addition layer, and output layer [3]. PNN is often being used for case of classifying due to able plotting every input into classification with quick processing and high accuracy more than any other ANN model[4]. Classic PNN is using euclidean distance and every class for classification is weighted with the same number of value [5]. Weighted Probabilistic Neural Network (WPNN) include weighting factors between the pattern layer and the summation layer of the conventional PNNs. The results show that modified PNN has improving the accuracy of classification [6].

2. Methodology

2.1. Analysis of Financial Statements
In order to learn that a company's financial health, whether in a good state level or the otherwise, financial report will be required. It can be achieved by comparing some financial ratios of the company itself from its income report and balance sheet report. With those comparison, we can learn its profitabilitiy, liquidity and solvency in certain times.

Based on the Decree of the Ministry of Finance of the Republic of Indonesia No. 740/KMK.00/1989 and its attachment regarding the weight of performance that has been revised in No. 826/KMK.013/1992, judging the company performance comes into 3 ratios, the first one is profitability ratio with 75%, then liquidity ratio with 12.5% proportion, and lastly solvency ratio with 12.5% proportion.

2.2. Determination of The Company's Financial Health Level
Profitability ratio, liquidity ratio, and solvency ratio can be used to determine the company's financial health level, which based on Decree of the Ministry of Finance of the Republic of Indonesia No. 826/KMK.013/1992. The calculation result will deterimine the company's financial health levels, that are:

a. Very good is when the company earned more than 110 score for its performance in recent year
b. Good is when the company earned between more than 100 to and less than equal to 110 score for its performance in recent year
c. Bad is when the company earned between more than 90 and less than equal to 100 score for its performance in recent year
d. Very bad is when the company earned less than equal to 90 score for its performance in recent year

In this study, the determination of company's financial health level are divided into 2 levels, which is:
1. Good state is when a company that earned more than 100 score for its performance in recent year.
2. Bad state is when a company that earned less than equal to 100 score for its performance in recent year.

2.3. Probabilistic Neural Network
Probabilistic Neural Network (PNN) is a classification algorithm that uses probabilistic function, does not require large dataset in learning stage, and has an advantages that can overcome the problems that exist in Back-Propagation (BP), which is shorten the amount of time processing in training stage, and difficulties in designing network architecture [7]. However, the PNN's problem lies in determining the smoothing parameter, which is no concrete way to define it, thus the value was obtained by trial and error until acquiring satisfaction result.

The PNN is inspired from Bayesian classification and classical estimators for probability density functions.These operations take placeat four layers of the PNN. The first layer is an input layer, The number of neurons in this layer is equal to the number of parameters (variables) needed to describe the

separable form input [8].The second layer contains nodes with activation functions. These nodes are called pattern units and are fully connected to the input nodes. There are as many pattern units as training samples. The feature values associated with each pattern unit are used as connection weights of that unit to the input layer nodes. The activation function that will be used is radial basis function [9], the formula is written below:

$$radbas(n) = e^{-n^2}, \text{ where n} = \text{neuron.}$$

The third layer contains summation units needed to complete the probability estimation. There are as many summation units as classes. Each summation unit receives input only from those pattern units belonging to its respective class. The last layer is the decision layer. The maximum of the summation node outputs is found in this layer. There are as many units as classes, each receiving input fromall summation units. Through a simple comparison, only the unit corresponding to the class with the highest summation unit value produces an output of one, while others generate a value of zero there by indicating the classification decision for the input vector [6].
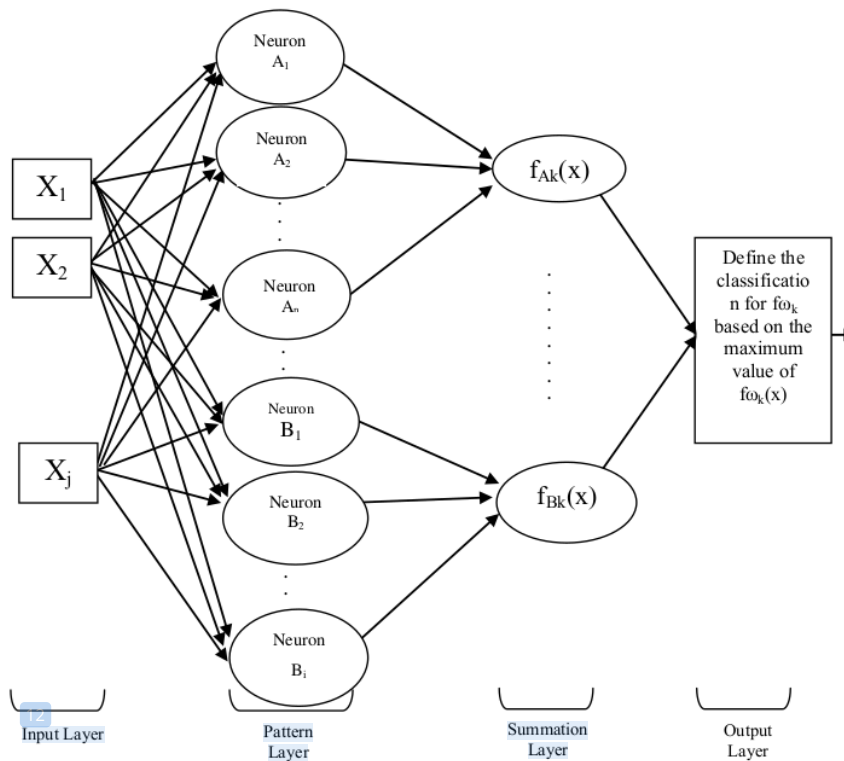

Figure 1. Structure of Probabilistic Neural Network Model

2.4. Weighted Probabilistic Neural Network

Modified PNN, or WPNNs include weighting factors between the pattern layer and the summation layer of the conventional PNNs. The architecture of the WPNNs is shown in figure 2. TheWPNNs are three-layer feed forward networks including an input layer, a pattern layer and a summation layer [5]. This weighted value will getting higher for patterns with high class separation levels, and vice versa.
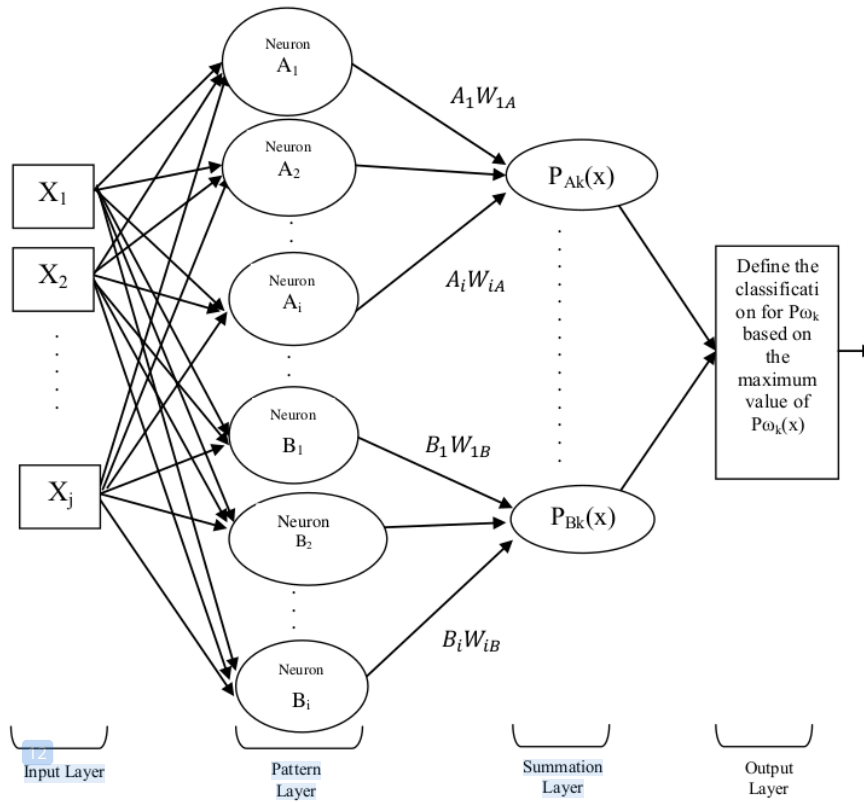
Figure 2. Structure of Weighted Probabilistic Neural Network Model

2.5. Classification Performance

According to Prasetyo [11], in order to measure the performance of a classification, commonly is by using confusion matrix. Table 1 is a confusion matrix 2×2 that classifies binary problems. Every cell f_{ij} in the matrix, states the number record (data) of the class i which predicted to join the class j. For example, cell f_{11} is the amount of data in class A correctly mapped to the class A, and cell f_{12} is the data in class A which wrongly mapped to class B.

Table 1. Confusion Matrix 2×2

f_{ij}		Predicted Class(j)	
		Class = A	Class = B
Original Class(i)	Class = A	f_{11}	f_{12}
	Class = B	f_{21}	f_{22}

4

Based on the contents of confusion matrix, to find the amount of data of each class that predicted correctly, is by using this formula $(f_{11} + f_{22})$, and incorrectly classified data, that is $(f_{12} + f_{21})$. By knowing the number of these values,the accuracy of the predictions and the rate of error of prediction will be able to conducted:

1. Accuracy $\qquad = \dfrac{f_{11}+f_{22}}{f_{11}+f_{12}+f_{21}+f_{22}}$

2. Apparent Error Rate $\qquad = \dfrac{f_{12}+f_{21}}{f_{11}+f_{12}+f_{21}+f_{22}}$

3. WPNN's Algorithm

Let **V** as training dataset. WPNN's testing algorithm is written below [9]:

1. Define **X** as matrix input with the size: $k \times j$

$$X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1j} \\ X_{21} & X_{22} & \cdots & X_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ X_{k1} & X_{k2} & \cdots & X_{kj} \end{bmatrix}$$

   X is the input data matrix with X_{kj}, where k represent as a number of data observation and j represent as a number of variable.

2. Set **T** as the target vector corresponding to the matrix **V.**

3. Initiation: $\qquad$ Weight for input layer: **V**

   $\qquad$ Weight for bias input layer: $b = \dfrac{\sqrt{-\ln(0.5)}}{spread}$, where $0 < \text{spread} \leq 1$

   $\qquad$ Weight for output layer: **W=T**

set $V = \begin{bmatrix} V_{11} & V_{12} & \cdots & V_{1j} \\ V_{21} & V_{22} & \cdots & V_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ V_{i1} & V_{i2} & \cdots & V_{ij} \end{bmatrix}$; $W = \begin{bmatrix} W_{1h} \\ W_{2h} \\ \vdots \\ W_{ih} \end{bmatrix}$

   Where **V** represent as matrix for training data where i represent as a number of observation and j represent as a number of variable. **W** is output matrix, where i represent as a number of observation and h represent as a number of class.

4. For k=1, for all data to be tested :

   a. Determined distance X_k to V_i, i=1, 2, ..., n as follows:

   $$D_{ki} = \sqrt{\sum_{j=1}^{p} (X_{kj} - V_{ij})^2}$$

   b. Determined the first activation function that is $a1_{ki}$ as follows:

   $\qquad a1_{ki} = e^{-(b_i D_{ki})^2}$; with i=1, 2, ..., n.

   c. Calculated mahalanobis distance for each data to be tested as follows:

   $\qquad d_i^{\,2} = [\vec{x_i} - \vec{m}]^T S^{-1} [\vec{x_i} - \vec{m}]$;

   $\qquad$ with i=1, 2, ..., n.

   $\qquad\qquad \vec{x}$= Vector with jcomponent

   $\qquad\qquad \vec{m}$= Vector mean – average of the independent variable with j component

   $\qquad\qquad S^{-1}$= Invers of variance matrix of the data being tested

   d. Determined second activation function namely $a2_{kh}$ as follows:

$a2_{kh} = \sum_{i=1}^{n}(a1_{ki}W_{ih})$; with W_{ih}= multiplication between mahalanobis distance to the vector target **T** corresponding to h=1, 2, …, r.

  e. Determined **Z** such that a2$_{kz}$ = max(a2$_{kh}$| h=1, 2, …, r).

  f. Determined **Z** as the class of X$_k$.

4. Variables

In this study, 140 Companies's Annual Financial report from official website of Indonesia Stock Exchange was used for classification of company performance. There are 7 variables that have been obtained from the data, which consist of 6 independent variables and 1 dependent category variable. Those variables are:

1. Dependent Variable

   The dependent variable in this study is about financial status of company that divided into 2 categories, which is Good State or Bad State

2. Independent Variable

   1. Working Capital to Total Asset Ratio (X_1)
   2. Earning Before Taxes Ratio (X_2)
   3. Total Asset Turnover Ratio (X_3)
   4. Debt to Asset Ratio (X_4)
   5. Long-term Debt to Asset Ratio (X_5)
   6. Return on Investment Ratio (X_6)

5. Result and Discussion

5.1. Financial Health Status

The description of the data shows the financial health status from 140 companies in Indonesia. Tabel 2 shows that only 35% or 49 companies was diagnosed in good state, while the rest, 65% or 91 companies in bad state. That's why, a model that capable to classfying the status with high accuracy is necessary.

Table 2. Description of Company's Financial Status

Company's Performance	The number of companies	Percentage
Good	49	35%
Bad	91	65%
Total	140	100%

5.2. Classification of Financial Health Status Using WPNN

Classification using WPNN was carried out for determining the optimum parameters, which is obtaining the highest classification rate. In other words, obtaining the minimum error by using Apparent Error Rate (APER) method.In the process of classfying with WPNN, the dataset was divided into 2 dataset, which is for training sample and testing sample. After several trial, for the optimum classification rate, in this dataset that consist 140 data observation, 90% of data (127 observation), was included into training sample, while the rest or 10% (13 observation) included into testing sample proportionally corresponding to the number of classification class.

  The next step is to defining the value of smoothing parameter (spread). In order to obtain the optimum value of spread, is by entering a series of number with interval from 0<spread≤1, in addition of holdout method in the process to increase the accuracy of classification, then repeating the calculation until a satisfying result was earned.The results will appear as shown in Table 3 below.

Table 3. WPNN's Classification Rate

spread	50:50		60:40		70:30		80:20		90:10	
	CR Train	CR Test	CR Train	CR Test	CR Train	CR Test	CR Train	CR Test	CR Train	CR Test
0.01	1.000	0.8000	1.000	0.8545	1.000	0.8537	1.000	0.9630	1.000	1.000
0.05	1.000	0.8286	1.000	0.8364	1.000	0.8293	1.000	0.8519	1.000	1.000
0.10	0.9857	0.8286	1.000	0.8545	1.000	0.8293	1.000	0.9259	0.9921	0.9231
0.20	0.9429	0.7143	0.9412	0.8000	0.9394	0.7317	1.000	0.8148	0.9134	0.7692
0.30	0.8857	0.8571	0.9294	0.7636	0.9293	0.7805	0.9204	0.7037	0.8898	0.8462
0.40	0.8714	0.8571	0.9059	0.7091	0.8788	0.8293	0.9115	0.8148	0.8740	0.6923
0.50	0.8286	0.8286	0.8824	0.6727	0.8687	0.7317	0.8673	0.7037	0.8661	0.6154
0.60	0.8571	0.6857	0.8824	0.6727	0.8687	0.7805	0.8673	0.7407	0.8425	0.6923
0.70	0.8714	0.7857	0.8588	0.6909	0.8788	0.6341	0.8584	0.7037	0.8425	0.7692
0.80	0.9000	0.7286	0.8588	0.8000	0.8687	0.6829	0.8496	0.8148	0.8346	0.7692
0.90	0.8286	0.7714	0.8588	0.7273	0.8485	0.7561	0.8584	0.6296	0.8268	0.7692
1.00	0.8714	0.7286	0.8588	0.6727	0.8586	0.6585	0.8407	0.8148	0.8268	0.7692

In Table 3 it appears 2 values that have the highest classification rate, which is 0.01 and 0.05. The classification rate for both value were 100% either in training stage nor testing stage.
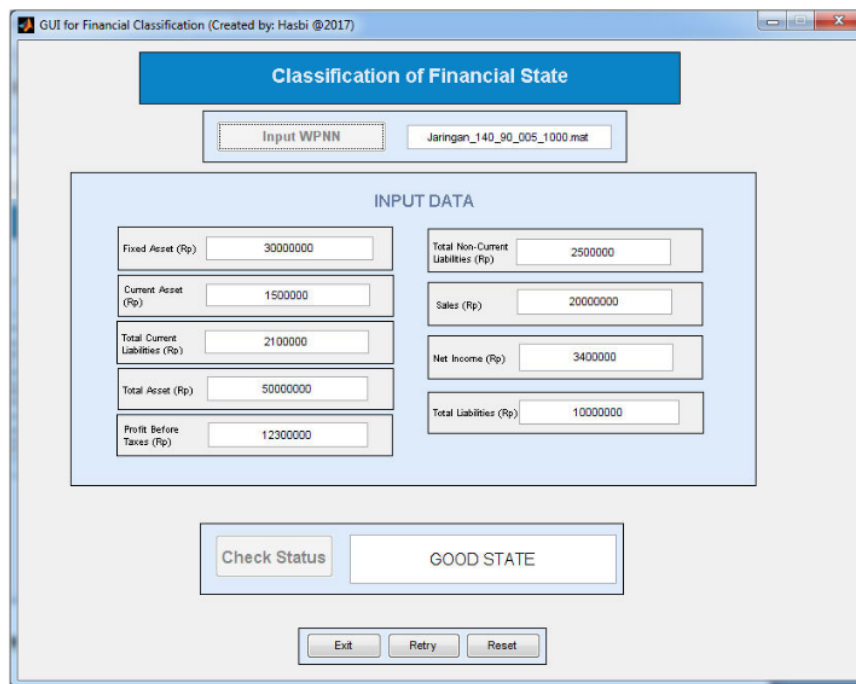

Figure 3. Display of GUI for Financial Classification

5.3. Predicting Classification of Company Performance Status

Having known that in this case, the best classification is obtained with data percentage of training sample and testing sample in 90:10, with accuracy evaluation using holdoutmethod, and smoothing parameter (spread) value of 0.01 and 0.05. Then in the examples below using that network with spread value of 0.05, the next step is testing the network itself when series of financial condition had been set.In this example, predicting the performance of company with fixed assets amounting to Rp 30,000,000.00, current assets amounting to Rp 1,500,000.00, total non-current (long-term) liabilites of Rp 2,500,000.00, total assets of Rp 50,000,000.00, profit before tax of Rp 12,300,000.00, total current (short-term) liabilites of Rp 21,000,000.00, total liabilites of Rp 10,000,000.00, sales of Rp 20,000,000.00 and net income of Rp 3,400,000.00.

In Figure 3 it appears that in this case wpnn's network predicted the company as having a good state in financial term, which means the performance of the company itself is in the good state as well.

6. Conclusion

WPNN is able to provide a high degree of accuracy in the classification of the company's financial performance status based on Working Capital to Total Asset Ratio, Earning Before Taxes Ratio, Total Asset Turnover Ratio, Debt to Asset Rati, Long-term Debt to Asset Ratio, and Return on Investment Ratio. The results show that the company's performance modeling with the WPNN model has a very high accuracy that reaches 100%.

References

[1]   Harahap M 2009 Analisis Kritis Laporan Keuangan Jakarta: Raja Grafindo Persada
[2]   Fausset L 1994 Fundamental of Neural Network: Architecture, Algorithm, and Application New
          Jersey: Prentice-Hall
[3]   Mishra M, Jena AR and Das R 2013 A Probabilistic Neural Network Approach For
          Classification of Vehicle *International Journal of Application or Innovation in Engineering
          & Management* Vol. **2**, No. **7**: 367-371
[4]   Sofha E, Yasin H and Rahmawati R 2015 Klasifikasi Data Berat Bayi Lahir Menggunakan
          Probabilistic Neural Network dan Regresi Logistik (Studi Kasus di Rumah Sakit Islam
          Sultan Agung Semarang Tahun 2014) *Jurnal Gaussian* Vol. **4** No. **4**
[5]   Ramakrishnan S and Selvan S 2006 Classification of Brain Tissues Using Muliwavelet
          Transformation and Probabilistic Neural Network *International Journal of Simulation:
          Systems, Sciencd and Technology, A Publication of the United Kingdom Simulation Society*,
          7(9):9-25
[6]   Ramakrishnan S and El Emary IMM 2009 Comparative Study between Traditional and
          Modified Probabilistic Neural Networks *Telecommunication System* **40**:67-74. DOI
          10.1007/s11235-008-9138-5
[7]   Specht DF 1990 Probabilistic Neural Networks, *Neural Networks* Vol.**3**: 109-118
[8]   Lotfi, A., &Benyettou, A., 2014. A Reduced Probabilistic Neural Network for The
          Classification of Large Databases, *Turkish Journal of Electrical Engineering & Computer
          Sciences* **22**: 979 – 989
[9]   Kusumadewi S 2004 Klasifikasi Pola Menggunakan Jaringan Probabilistik *Seminar Nasional
          Aplikasi Teknologi Informasi* pp: J59-J64
[10]   Yasin H, Ispriyanti D and Rusgiyono A 2016 Klasifikasi Risiko Kredit Menggunakan Weighted
          Probabilistic Neural Network (WPNN), *Prosiding Seminar Nasional Variansi 2016*, ISBN
          978-602-99837-2-2: 58-69
[11]   Prasetyo E 2014 Data Mining: Konsep dan Aplikasi Menggunakan Matlab.Yogyakarta: Andi
          Publisher

2018 Yasin_2018_J._Phys.%3A_Conf._Ser._1025_012095

Internet Source 1%

6 ijssst.info
Internet Source 1%

7 Shuai Tao, Xiaowei Zhang, Huaying Cai, Zeping Lv, Caiyou Hu, Haiqun Xie. "Gait based biometric personal authentication by using MEMS inertial sensors", Journal of Ambient Intelligence and Humanized Computing, 2018
Publication 1%

8 www.idosi.org
Internet Source 1%

9 Wenhui Hou, Ye Wei, Jie Guo, Yi Jin, Chang'an Zhu. "Automatic Detection of Welding Defects using Deep Neural Network", Journal of Physics: Conference Series, 2018
Publication 1%

10 www.ejournal-s1.undip.ac.id
Internet Source 1%

11 Submitted to University of Malaya
Student Paper 1%

12 "Soft Computing Applications", Springer Nature America, Inc, 2016
Publication 1%

13 Tatenda Duncan Kavu, Tinotenda Godknows Nyamandi, Alleta Chirinda, Talent T. Rugube, 1%

Kudzai Zishumba. "Human Action Prediction Using Sentiment Analysis on Social Networks", International Journal of ICT Research in Africa and the Middle East, 2017
Publication

14 Submitted to HELP UNIVERSITY
Student Paper
1 %

15 jcc.undip.ac.id
Internet Source
1 %

16 krishnasamy, Umamaheswari, and Devarajan Nanjundappan. "Hybrid weighted probabilistic neural network and biogeography based optimization for dynamic economic dispatch of integrated multiple-fuel and wind power plants", International Journal of Electrical Power & Energy Systems, 2016.
Publication
1 %

17 www.kmice.cms.net.my
Internet Source
<1 %

18 www.rroij.com
Internet Source
<1 %

19 Cornelius Glackin, Liam Maguire, Ronan McIvor, Paul Humphreys, Pawel Herman. "A comparison of fuzzy strategies for corporate acquisition analysis", Fuzzy Sets and Systems, 2007
Publication
<1 %

20   S. Ramakrishnan. "Classification brain MR images through a fuzzy multiwavelets based GMM and probabilistic neural networks", Telecommunication Systems, 02/18/2010
Publication

<1 %

21   jise.ir
Internet Source

<1 %

22   Submitted to Universitas Sumatera Utara
Student Paper

<1 %

23   Taufik Djatna, Ida Bagus Dharma Yoga Santosa. "A real time mission-critical business intelligence for development of mixture composition on aromatherapy product based on customer personality type", 2014 6th International Conference on Information Technology and Electrical Engineering (ICITEE), 2014
Publication

<1 %

24   dspace.lboro.ac.uk
Internet Source

<1 %

25   Lecture Notes in Computer Science, 2015.
Publication

<1 %

26   Stone, . "Credit Card Receivables", The Securitization Markets Handbook Structures and Dynamics of Mortgage- and Asset-backed Securities, 2012.
Publication

<1 %

27 Lecture Notes in Computer Science, 2005.
Publication

<1 %

Exclude quotes   Off   Exclude matches   Off
Exclude bibliography Off

2018 Yasin_2018_J._Phys.%3A_Conf._Ser._1025_012095