International Journal of Computer Applications Technology and Research
Volume 10–Issue 07, 192-197, 2021, ISSN:-2319–8656

Analysis of the Compatibility of Student Satisfaction Index with Course Value Using Significance Correlation and Backpropagation Approach

Iswahyudi Widodo
Electrical Engineering Department,
Faculty of Engineering,
Brawijaya University.
Malang, East Java, Indonesia

Sholeh Hadi Pramono
Electrical Engineering Department,
Faculty of Engineering,
Brawijaya University.
Malang, East Java, Indonesia

Muhammad Aswin
Electrical Engineering Department,
Faculty of Engineering,
Brawijaya University.
Malang, East Java, Indonesia

Abstract: The implementation of the measurement of customer satisfaction is the measurement of service satisfaction of lecturers to students to lecturers carried out by the State Islamic University of Maulana Malik Ibrahim Malang during the pandemic using the IKM Online application. The competencies measured are pedagogic competence, professional competence, personality competence, and social competence. The measurement of the correlation of significance to the results of IKM Online compared with student scores as learning outputs is carried out to determine the relationship between the two variables. The coefficient of determination is also added to the measurement to determine the strength of the influence of one variable on another variable. The results show that the correlation between the 2 variables is strong and the subject value variable has an effect of 55.9504% on the IKM outcome variable. The data classification process using the backpropagation algorithm is used to strengthen the relationship between the two variables. The result is the 2-5-1 model with RMSE 0.028 being the best model and the strong data class being the first class. From this study, measurements are needed to determine other factors that affect the assessment of student satisfaction to lecturers and evaluation of survey instruments.

Keywords: satisfaction survey, IKM online, correlation, significance, coefficient of determination, backpropagation

1. INTRODUCTION

Customer satisfaction is one of the determining indicators to determine whether the business processes carried out by a business unit have been successful or not. The more satisfied with the goods or services provided to customers, the better the products presented to customers [1]. Therefore, it is not surprising that customer satisfaction always gets special attention when formulating a business strategy [2].

The measurement of service satisfaction is not without questions. The results of the evaluation conducted by the Quality Assurance Institute (LPM) of State Islamic University of Maulana Malik Ibrahim Malang found that there were doubts about the results of the satisfaction survey. The doubts are related to whether the survey results are in line with the performance of the lecturers who are being assessed [3]. Whether the survey respondents gave an assessment of the value they received. Then whether the assessment is carried out randomly without looking at the aspects measured in the survey, given the pandemic conditions resulting in direct interaction of lecturers and students in the learning process cannot be carried out. Then whether it is in accordance with the aspects measured.

To answer this question, it is necessary to carry out several analyzes, one of which can be done is to analyze the relationship between the results of measuring student satisfaction with lecturer services and the value of the courses obtained by students as the output of learning (learning output) with the correlation significance method. [4]. Comparing the behavior pattern approach with this student score variable, a value approach will be obtained which will describe whether or not there is a correlation between the two. Tracing the pattern of respondent behavior that can be tracked through filling out the survey. From this classification, a behavioral model of survey respondents will be obtained and to compare it with the measured correlation results.

The search for data classification is by measuring variables such as the Student Satisfaction Index (IKM) which is measured using the IKM Online application, the value of students as learning outcomes with the variable of the course lecturer. The method used to classify these variables is an artificial neural network approach with the Backpropagation method. Although it has several weaknesses, such as training results that are not constant and it is not known in detail how the prediction results are obtained, because this method cannot provide information about the most influential weights among the input patterns, this method also has advantages. The advantages of this method are that it is able to formulate the experience and knowledge of forecasters, and is very flexible in changing forecasting rules [5].

The object in this study is the result of a survey of lecturer service satisfaction to students in the Covid-19 pandemic condition using the IKM Online application at UIN Maulana Malik Ibrahim Malang in the odd semester of 2020/2021. The purpose of this study is to measure the level of correlation between the implementation of the measurement of lecturer service satisfaction to students compared to learning outcomes in the form of academic values with the correlation significance method and artificial neural networks. This study is also intended to find the relationship between the two variables and perform a classification process on the measurement results. The results of these measurements are also to determine whether there is a correlation between the measurement of service satisfaction of lecturers and students with the learning process. With this research, it is expected to be able to answer the results of the evaluation findings and an improvement in the measurement of IKM and the quality of

learning at State Islamic University of Maulana Malik Ibrahim Malang in the future [6].

2. METHOD

The concept of this research method is as illustrated in Figure 1 below:

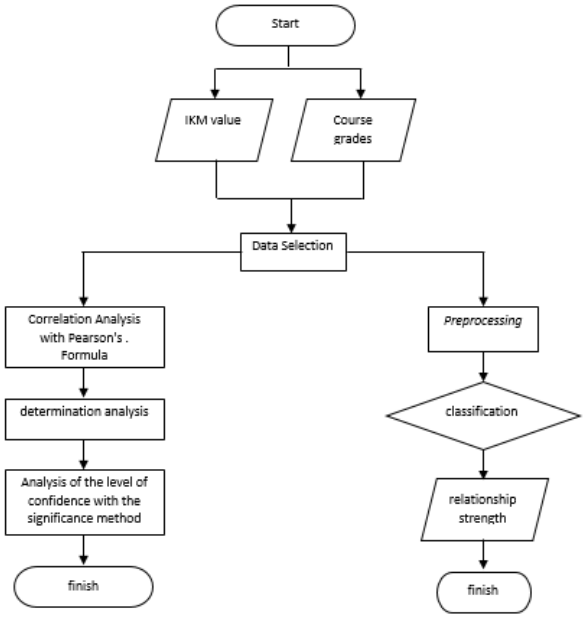

Figure 1. Research Method

2.1 Data Selection

Data selection is the initial stage in preparation before the data is analyzed [7]. In the data selection, the IKM result data is taken for the nim variable, the IKM value, the course code and the lecturer code. The student value data is taken for the variable nim, course value, course code and lecturer code. The results of this data selection are data tables for nim, ikm value, mk value, mk code, lecturer code.

2.2 Preprocessing

Preprocessing is the initial stage for classification analysis. At this stage the attributes and data classes are determined. The data is then normalized so that the normalized data is ready to be processed into the system.

2.3 Correlation Coefficient of Significance and Determination

This is the most widely used correlation analysis formula, which measures the strength of the 'linear' relationship between the raw data of the two variables, rather than their ranking [8]. This formula is a dimensionless coefficient, meaning that there are no data-related limitations to consider when performing an analysis with this formula, which is the reason why this coefficient is a suitable formula to use [9].

$$r = \frac{n\sum x_i\, y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i{}^2 - (\sum x_i)^2}.\sqrt{n\sum y_i{}^2 - (\sum y_i)^2}}$$

Where x is the IKM result variable and y is the subject value variable. the resulting measurement will be in numerical form generated by all correlation coefficients, including Spearman's Rank and Pearson's Coefficient: $-1 < r < +1$.

The method described above should be used according to whether there are parameters associated with the data collected or not. Two terms to pay attention to are:

a. Parametric: (Pearson's coefficient) Namely where data must be handled in terms of population parameters or probability distributions. Usually used with quantitative data already defined in the parameter.
b. Nonparametric: (Spearman's Rank) That is where no assumptions can be made about the probability distribution. Usually used with qualitative data, but may be used with quantitative data if Spearman's Rank proves inadequate.

The correlation coefficient, *r*, provides only a measure of the strength and direction of a linear relationship between two variables. However, it does not provide information about the proportion of variation (variation) of the dependent variable (Y) that can be explained or caused by a linear relationship with the value of the independent variable (X). The value of r cannot be directly compared, for example we cannot say that the value of $r = 0.8$ is twice the value of $r = 0.4$. However, the squared value of *r* can accurately measure the ratio/proportion, and this statistical value is called the Coefficient of Determination, r^2. The coefficient of determination can be defined as a value that expresses the proportion of variance Y that can be explained/explained by a linear relationship between variables X and Y.

The level of significance, also known as alpha or , is a measure of the strength of the evidence that must be present in a study sample before the researcher rejects the null hypothesis and concludes that the effect is statistically significant. Researchers determine the level of significance before conducting the experiment. The level of significance is the probability of rejecting the null hypothesis even though the hypothesis is true. Significance can be measured by the equation:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}.$$

where n is the number of samples and r is the Pearson's correlation.

2.4 Artificial Neural Network

An artificial neural network (ANN) or also called a Neural Network is a network of a collection of small-level processing units that are described based on a biological nervous system. An artificial neural network can change its structure in order to solve a problem based on information obtained both internally and externally through the system, so it is also called an adaptive network because of its adaptive capabilities [10]. This Artificial Neural Network is determined by 3 things, namely:

1. The pattern of relationships between neurons or so-called network architecture
2. Algorithm which is a method for determining the weight of the link or link or can also be called the training process
3. Activation function

An artificial neural network is a mathematical model that defines the function f: $X \rightarrow Y$. The network in this system is an interconnection of neurons that are in different layers. The layers in the ANN are divided into 3, namely:

1. Input layer or input layer.
2. Hidden layer or hidden layer.
3. Output layer or outer layer.

Mathematically, a neuron is a function that receives input from the previous layer $g_i(x)$ (i layer). This function generally processes a vector and then converts it to a scalar value, as the equation $f(x) = K(\Sigma i\ w_i\ g_i\ (x))$, where K is a special function

which is often called the activation function and w is the load or weight.
In classification, there are several methods that can be used to classify them [9]:

1. Artificial Neural Network (ANN)
2. Naive Bayes
3. Support Vector Machine (SVM)
4. Decission Tree
5. Fuzzy

Each method has advantages and disadvantages and also different specifications for each use, one of which is a data object. The use of methods with different object types will result in different optimizations.

2.5 Backpropagation

Backpropagation is one of the learning algorithms in Artificial Neural Networks [11]. The learning mechanism in backpropagation is done by adjusting the ANN weights in a backward direction based on the error value in the learning process. Backpropagation consists of three layers, namely the input layer, hidden layer and output layer.

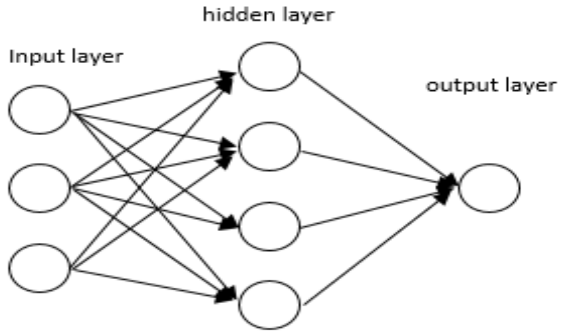

Figure 2. Backpropagation algorithm architecture

2.5.1 Advantages and Disadvantages of Backpropagation Algorithm

This backpropagation algorithm also has advantages and disadvantages. Among the advantages of this algorithm are [12]:

a. Can be implemented in solving a problem related to identification, prediction, pattern recognition and others.
b. Have the ability to learn and be immune to mistakes, so as to form a system that is resilient and can work consistently.
c. In the training process, the network can get a balance so that it is able to provide the correct response to input patterns such as the patterns used during training.

The disadvantages of this algorithm are:

a. The learning rate parameter can vary according to the condition of the error change in each iteration.
b. It takes a long time in the learning process to arrive at the convergent level.
c. The calculation of the weight change of the algorithm is able to cause a local minimum, thus causing instability.

2.5.2 Backpropagation Training

Backpropagation algorithm training is carried out to get the desired result estimate [13]. Some steps in backpropagation algorithm training are as follows [14]:

Step 1: Initialize the weights with the smallest random number
Step 2: As long as the stop condition is false, do steps 3 – 9 feed forward process
Step 3: Each input unit (x_i, i = 1,…, n); receive input signal and forward the signal to all hidden layers.
Step 4: Each hidden unit (z_j, j = 1,…, p); add up the weights of the signals input signal (including bias)

$$z\_in_j = v_{0j} + \sum_{i=1}^{n} (x_i . v_{ij})$$

then calculate the output signal from the hidden layer using the activation function count $z_j = f(z\_in_j)$. This output signal is then sent to all layers in the output layer.
Step 5: Each output (y_k, k = 1,…, m); summing the input signal weights:

$$Y\_in_k = w_{0j} + \sum_{k=1}^{p} (z_j . v_{jk})$$

then calculate the output signal from the hidden layer using the activation function count $Yj = f(Y\_in_k)$. This output signal is then sent to all layers in the output layer.

Process feedback (backward)
Step 6: Each unit of output (y_k, k = 1,…, m); receive a pattern according to the input training, to calculate the error between the target and the output generated network;

$$\delta_k = (t_k - y_k) f'(y_{in_k})$$

The factor δ_k is used to calculate the error correction w_jk which will later will be used to fix w_{jk}

$$\Delta w_{jk} = \alpha \delta_k z_j$$

Also count w_ok to fix w_{ok}

$$\Delta w_{ok} = \alpha \delta_k$$

Step 7: Each hidden unit (Zj, j = 1,…, p); summing input delta from previous step

$$\delta\_in_j = \sum_{k=1}^{m} \delta_k w_{jk}$$

Then calculate the weight correction with $v_{ij}=\alpha\delta_j x_i$ and calculate the correction bias using $v_{oj}=\alpha\delta_j$.
Step 8: Each output unit (Yk, k = 1,…, m); fix the weight and bias of hidden layer (j = 0,…, p) where w_{jk} (new)=w_{jk} (old) + w_{jk}
Step 9: The test stops when an error is found and training is stopped.

3. RESULT AND DISCUSSION

3.1 Data Selection

The data selection process is carried out after data collection. The data taken is data on the results of online IKM in the odd semester of 2020/2021 and data on the value of students in the odd semester of 2020/2021 taken from the online SIAKAD application. From these data, information is obtained in the IKM data there are 794,535 rows of student entries. As for the student value data obtained as many as 135,536 rows of data. The two data above are then queryed using PHP and MySql applications as shown below.

id	nim	kode_mk	nilai_mk	kode_dosen	nilai_ikm
1	17520099	952305	3	51060	4,0
2	17520074	952315	8	52026	4,5
3	17520014	952405	5	51053	4,0
4	17520046	952405	1	51053	5,0
5	16510086	1051215	7	51001	5,0
6	16510120	1051215	3	51001	4,0
7	16510140	1051215	7	51001	5,0
8	17510103	1051215	7	51001	4,0
9	17510118	1051215	9	51001	4,5
10	17510224	1051215	1	51001	3,0
11	15510130	1051306	8	51086	5,0
12	17510027	1051308	1	51077	3,0
13	17510027	1051314	1	51024	3,0
14	16510075	1051314	6	51024	3,1
15	16510086	1051314	3	51024	5,0
16	16510120	1051314	1	51024	3,5
17	17510145	1051317	9	51082	4,0
18	17510224	1051317	9	51082	3,5
19	16510086	1051317	9	51082	5,0
20	17510081	1051318	8	51006	5,0
21	17510145	1051318	8	51006	4,0
22	17510143	1051402	7	51012	2,6
23	17510170	1051402	9	51012	3,0
24	17510196	1051402	5	51012	3,0
25	16510086	1051402	6	51012	5,0
26	16510086	1051406	3	51014	5,0
27	14660064	1266503	6	66010	3,1
28	15660073	1266503	1	66007	4,0
29	16660020	1266503	8	66007	4,8
30	16660122	1266503	1	66007	5,0
31	16660123	1266503	1	66010	4,0
32	16540050	1354213	8	51069	2,2
33	18540114	1354213	8	51069	4,8
34	18540115	1354213	8	51069	4,0
35	18540114	1354318	3	54007	4,7
36	18540065	1354404	8	51061	4,0
37	18540056	1354405	8	51021	3,9
38	17540082	1354408	5	54014	5,0
39	14210058	1421223	8	21075	3,0
40	17210156	1421227	8	21057	3,0
41	17210190	1421412	8	21007	3,0
42	17210110	1421500	9	21010	3,0
43	18220066	1422233	5	20002	5,0

Figure 3. Data processing results

Processed data obtained through data selection were analyzed by correlation method with Pearson's formula. The correlation results from data processing using PHP and MySQL above are

r = 0.748

and the coefficient of determination is obtained

r^2 = 0.559504

The significance obtained with the equation below and processed with the script above is:

t=(r-√(n-2))/√(1- r^2)

=(0.748-√(58024-2))/√(1- 0.559504)

This calculation illustrates that the close relationship of these two variables is 0.748 or 74.8%. To determine the effect of one variable with other variables, it can be used to calculate the coefficient of determination. From the calculation using the coefficient of determination equation obtained r2 = 0.559504. This means that the variable value of the course has an influence relationship on the IKM outcome variable of 0.559504 or 55.9504%. This condition shows that the results of the IKM survey are only influenced by the variable value of 55, 9504%. The other 44.0496% are influenced by other variables.

3.2 Preprocessing

This process is a process to prepare data before classification. Merged data is normalized. Data is also processed to determine variables and classes for mapping. Normalization of data is done with PHP and MySql applications using the equation:

X'=0,8(X-b)/((a-b))+0,1

Where:

X' = normalized data

X = original data/data to be processed

a = maximum value of original data

b = minimum value of original data

The results of processing the data are partially illustrated in Figure 5 below.

id	nim	kode_mk	kode_dosen	nilai_mk	nilai_ikm
1	17520099	952305	51060	0,3	0,7
2	17520074	952315	52026	0,8	0,8
3	17520014	952405	51053	0,5	0,7
4	17520046	952405	51053	0,1	0,9
5	16510086	1051215	51001	0,7	0,9
6	16510120	1051215	51001	0,3	0,7
7	16510140	1051215	51001	0,7	0,9
8	17510103	1051215	51001	0,7	0,7
9	17510118	1051215	51001	0,9	0,8
10	17510224	1051215	51001	0,1	0,5
11	15510130	1051306	51086	0,8	0,9
12	17510027	1051308	51077	0,1	0,5
13	17510027	1051314	51024	0,1	0,5
14	16510075	1051314	51024	0,6	0,514285714
15	16510086	1051314	51024	0,3	0,9
16	16510120	1051314	51024	0,1	0,6
17	17510145	1051317	51082	0,9	0,7
18	17510224	1051317	51082	0,9	0,608571429
19	16510086	1051317	51082	0,9	0,9
20	17510081	1051318	51006	0,8	0,9
21	17510145	1051318	51006	0,8	0,7
22	17510143	1051402	51012	0,7	0,42
23	17510170	1051402	51012	0,9	0,5
24	17510196	1051402	51012	0,5	0,5
25	16510086	1051402	51012	0,6	0,9
26	16510086	1051406	51014	0,3	0,9
27	14660064	1266503	66010	0,6	0,52
28	15660073	1266503	66007	0,1	0,7
29	16660020	1266503	66007	0,8	0,86
30	16660122	1266503	66007	0,1	0,9
31	16660123	1266503	66010	0,1	0,7
32	16540050	1354213	51069	0,8	0,34
33	18540114	1354213	51069	0,8	0,86
34	18540115	1354213	51069	0,8	0,7
35	18540114	1354318	54007	0,3	0,84
36	18540065	1354404	51061	0,8	0,7
37	18540056	1354405	51021	0,8	0,68
38	17540082	1354408	54014	0,5	0,9
39	14210058	1421223	21075	0,8	0,5
40	17210156	1421227	21057	0,8	0,5
41	17210190	1421412	21007	0,8	0,5
42	17210110	1421500	21010	0,9	0,5
43	18220066	1422233	20002	0,5	0,9

Figure 4. Data normalization results

This data will then be used as processing data and training data that will be entered into the RapidMiner application. The determination of attributes and data classes is done by division as in the table below.

Table 1. Table of ikm attribute distribution

Number	Ikm	Value	Class
1	1	1	5
2	1	2	4
3	1	3	3
4	1	4	2
5	1	5	1
6	2	1	2
7	2	2	1
8	2	3	2
9	2	4	3
10	2	5	4
11	3	1	3
12	3	2	2
13	3	3	1
14	3	4	2
15	3	5	3
16	4	1	4
17	4	2	3
18	4	3	2
19	4	4	1
20	4	5	2
21	5	1	5
22	5	2	4
23	5	3	3
24	5	4	2
25	5	5	1

System formation, learning and testing processes are shown in figure 6 below

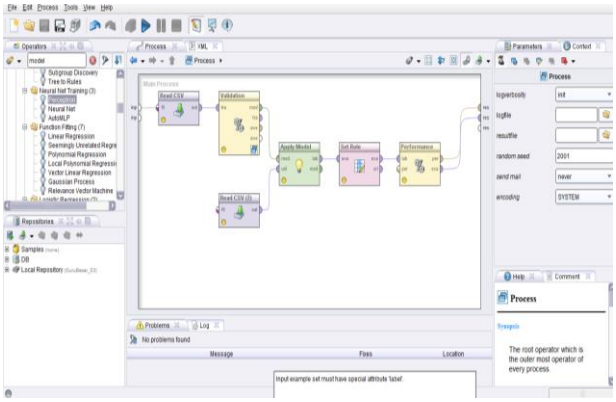

Figure 5. Establishment of the learning and testing process

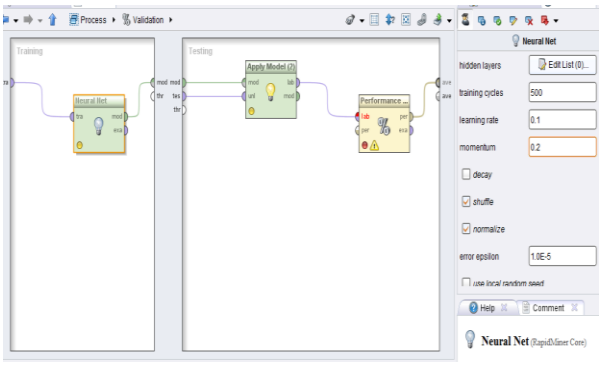

Figure 6. Backpropagation architecture on X-Validation operator

The parameters used in the backpropagation architecture to solve the above problem are:

Input	: 4
Training cycles	: 500
Learning rate	: 0.1
Momentum	: 0.2
Output	: 1
Error target	: 0.0001
Hidden	: 1 layer, 2 layer
Hidden 1 layer	: 2, 5, 8 neuron
Hidden 1 layer	: 2-5, 5-8 neuron

The above model is trained to choose the best architectural model as measured by the Root Mean Square Error (RMSE) parameter. The backpropagation model in the RapidMiner application can be created by clicking on the neural network (Figure 5.10) and selecting hidden layers in the parameters. The results of the implementation of the above model are as in Table 2 below:

Table 2. Root Mean Square Error (RMSE) training results with the Backpropagation algorithm

No	Model	RMSE
1	4-2-1	0.025 +/- 0.000
2	4-5-1	0.028 +/- 0.000
3	4-8-1	0.026 +/- 0.000
4	4-2-5-1	0.024 +/- 0.000
5	4-5-8-1	0.025 +/- 0.000

The results of processing using the RapidMiner application show that the backpropagation algorithm with the 4-5-1 model produces RMSE 0.028 with class division as Table 5.4 has the suitability of being the optimal model in measurement. In the class division, it also appears that the class with a. It's just that it should be noted that the RMSE value depends on the number of variables and the data that is processed. The more variables that are processed, the higher the value and the better model will also change. The result of a strong class status is more than any other status. This result is in accordance with the correlation approach. It's just that this classification also does not make the measurement of the value of the influence approach between variables visible in the classification. Another measurement method is needed that allows to measure the influence between variables.

4. CONCLUTION

From the research on the suitability analysis of the student satisfaction index with the value of the course using the significance correlation approach and backpropagation, the following results were obtained:

1. The results of the correlation measurement using the Pearson's correlation method produce a correlation value of r = 0.748 which means that the close relationship between the two variables is 0.748 or 74.8%.
2. The results of the correlation measurement using the coefficient of determination method produce a value of the coefficient of determination r2 = 0.559504. This means that the variable mk_value affects the result_ikm variable by 0.559504 or 55.9504%. This also means that there are 44.0496% factors from other variables that affect the measurement.
3. The results of the correlation measurement using the Pearson's correlation method yielded a significance value of t = 271.4732. This shows that the level of confidence in the measurement of the two variables above is very high.
4. The use of artificial neural network classification using the backpropagation algorithm with the help of the RapidMiner application to obtain optimal results of data processing using a 4-5-1 model with RMSE 0.0028 and classification with data classes according to the specified model. The class with strong status becomes the class that appears a lot in data processing.

From this research, it is hoped that further research will be carried out to determine other variables that influence the measurement of lecturer service satisfaction to students at State Islamic University of Maulana Malik Ibrahim Malang with the IKM Online application with student respondents. This determination is to test whether the instruments used for measurement in this pandemic condition are still reliable or whether improvements need to be made in describing the 4 competencies measured in the survey or need to be replaced. Evaluation and improvement in this measurement is expected to be able to provide an overview of service satisfaction from lecturers to students in learning to be closer to the truth.

5. REFERENCES

[1] Nagel, Pieter J.A.., Cilliers, Willem W.. 1990. Customer Satisfaction: A Comprehensive Approach.

[2] Rogado Gonzalez, Susana., Almos. 2014. Key Factors for Determining Student Satisfaction in Engineering: A Regression Study

[3] Lembaga Penjaminan Mutu. 2021. Laporan Audit Internal Indeks Kepuasan Mahasiswa terhadap Kinerja

Pembelajaran Dosen UIN Maulana Malik Ibrahim Malang.

[4] Huang, Ding-wei. 2016. Positive correlation between quality and quantity in academic.

[5] Widodo, A.P., Suhartono, Sarwoko, E.A., Firdaus, Z. 2017. Akurasi Model Prediksi Metode Backpropagation Menggunakan Kombinasi Hidden Neuron Dengan Alpha.

[6] Paolin, Allison. 2015. Enhancing Teaching Effectiveness and Student Learning Outcomes.

[7] Dewi, Sari. 2016. Komparasi 5 Metode Algoritma Klasifikasi Data Mining Pada Prediksi Keberhasilan Pemasaran Produk Layanan Perbankan.

[8] Asuero, A.G., Sayago, A., Gonz´alez, A.G. 2006. The Correlation Coefficient: An Overview. Critical Reviews in Analytical Chemistry

[9] Al-Sammaraie, N.A., Al-Mayali, Y.M.H., El-Ebiary, Y.A.B. 2018. Classification and diagnosis using backpropagation Artificial Neural Networks (ANN) algorithm.

[10] Wibawa, A.P., Purnama, Aji, M.G., Akbar, Fathony M., Dwiyanto, Andika F. 2018. Metode-metode Klasifikasi.

[11] Setti, Sunil., Wanto, A. 2018. Analysis of Backpropagation Algorithm in Predicting the Most Number of Internet Users in the World.

[12] Windarto, A.P., Nasution, D., Tambunan, A.W.F. 2020. Jaringan Syaraf Tiruan:Algoritma Prediksi dan Simulasi.

[13] Maharani, Warih. 2009. Klasifikasi Data Menggunakan JST Backpropagation Momentum dengan Adaptif Learning Rate.

[14] Alan, S.B., Pramono, Sholeh Hadi., Suyono, H. 2018. Konsolidasi Beban Kerja Kluster Web server Dinamis dengan Pendekatan Backpropagation Neural Network.