5

by Al Am

Submission date: 14-Jun-2021 03:28PM (UTC+0700)
Submission ID: 1606178546
File name: Walid_2019_J._Phys.__Conf._Ser._1321_032110.pdf (835.79K)
Word count: 3420
Character count: 17302

Naïve Bayesian classifier algorithm and neural network time series for identification of lecturer publications in realizing internationalization of Universitas Negeri Semarang

Naïve Bayesian classifier algorithm and neural network time series for identification of lecturer publications in realizing internationalization of Universitas Negeri Semarang

Walid[1,*] and Alamsyah[2]

[1]Mathematic Department, Universitas Negeri Semarang, Indonesia
[2]Computer Science Department, Universitas Negeri Semarang, Indonesia

*Corresponding author : alikuwalid87@yahoo.com

Abstract. The Naïve Bayes Classifier is a classification method rooted in the Bayes theorem. The main characteristic of the Naïve Bayes Classifier is the very strong (naïf) assumption of the independence of each condition/ event. If X is the vector that enters the feature and Y is the class label, Naïve Bayes is written with P (X | Y). This notation means that the probability of a class Y label is obtained after the X features are observed. This notation is also called the posterior probability for Y, while P (Y) is called the prior probability Y. One of the important factors included in the UNNES internationalization assessment is the results of lecturer publications both at the national and international levels. The artificial neural network as one of the information processing systems designed by imitating the workings of the human brain in completing a problem by doing the learning process through changes in synaptic weight can be used to analyze the productivity of the performance of UNNES lecturers and staff in realizing the year of reputation proclaimed in 2017 Research related to this theme has been carried out in 2017 and shows that the factors of knowledge and behavior of lecturers and education staff have the most influence on the achievement of the year of reputation at FMIPA UNNES in 2017.

1. Introduction

Classification uses the probability method in predicting opportunities for future events based on past experience or data often experts using Naïve Bayes from this branch of science born the Bayes Theorem concept Artificial neural network applications have developed rapidly, including for forecasting purposes. The main advantage of artificial neural networks is the ability of parallel computing by learning from the patterns taught. Artificial neural networks can be trained to study and analyze past data patterns and try to find a formula or function that will link past data patterns with the desired output at this time. Artificial neural network models are determined by network architecture and training algorithms.

There are various kinds and types of stationary time series models known in the world of statistics including the Autoregressive (AR) model, the Moving Average (MA) model, and the combined or mixed model of both the ARMA model. The time series model that is not stationary is often called the ARIMA (Autoregressive Integrated Moving Average) model. Semarang State University (UNNES) has established 2018 to be one of the internationalization years, one of which is in the field of research and service. One of the targets for international shielding is to increase the quality and quantity of lecturers' scientific publications, both at national and international levels.

The development of science is so dynamic and complex, one of which is the development of artificial neural networks (JSN). How to get a good JSN model, that is through determining the right combination between input variables and units in the hidden layer, is a central topic in some JSN literature that has been widely discussed in many articles and many books [1].

Nowadays there are many studies that use forecasting models and electricity using neural networks in various fields, including in the fields of labeling Taylor according to [2], [3], [4],[5] and [6}. The development of research is related to the RNN model in the ARFIMA process of long memory data as developed by [4]. While research that develops the effect of data outliers can be seen [8]. Developed recurrent neural networks for forecasting time series with long memory patterns as explained in [9].

Based on the explanation above, a scientific study will be developed on the use of the Naïve Bayes classifier time series neural network algorithm (TSNN) in the identification of lecturer *publications* at FMIPA UNNES in realizing internationalization programs in the field of research and service. Based on this identification, what factors will influence the development of lecturer publications both at the national and international levels as an effort to realize the internationalization of the UNNES.

Based on the description above, the focus of the problems in this study are as follows. (1) What is the model of identification of lecturer publications in realizing internationalization of UNNES using the Naïve Bayes classifier algorithm time series neural network? (2) What is the result of identifying lecturer publications in realizing internationalization of UNNES using the Naïve Bayes classifier algorithm time series neural network?

1.1. Naïve Bayes theorem and Neural Network Time Series Models

A related explanation about the Naïve Bayes model on its use in the field of decision, then in the calculation of opportunities provided that the decision class is correct, considering the vector information object [10]. Naive Bayes is a simple classification algorithm where each attribute is independent and allows to contribute to the final decision. The assumption of this algorithm is that object attributes are independent, so the opportunities that occur are calculated based on the number of frequencies.

The basis of the Naïve Bayes theorem used in programming is the Bayes formula, which is as follows [10]

$$P(H|E) = \frac{P(E|H).P(H)}{P(E)} \tag{2.1}$$

1.2. Classification Tree (Decision Tree)

The decision tree model is a simple representation of classification techniques that are often used in finite classes, in this class the internal nodes and also root nodes are marked using attribute names, while the ribs are labeled with attribute values, while leaf nodes are marked using classes that varies

Often the decision tree model is called the classification tree, this model is one of the most popular data mining processes among researchers because often this model is used for purposes of solving problems in daily life. One of the advantages of this model is that the accuracy in the data classification process by passing the data record to the leaf node from the decision tree uses the attribute variable values and places the target node's leaf value into the data record.

1.3. Classification with Naïve Bayes Classifier

One classification model that has a close relationship with the Bayes theorem concept is the Naïve Bayes Classifier model. This model has special characteristics, namely with assumptions that are so strong against the nature of independence in all circumstances Suppose X is a vector that has a feature value and Y is a class label, then the Naïve Bayes model can be expressed in form P(X|Y) as explained in [12].

The Naïve Bayes formulation for classification is:

$$\mathrm{P(Y|X)} = \frac{\mathrm{P(Y)}\prod_{i=1}^{q}\mathrm{P(X_i|Y)}}{\mathrm{P(X)}} \tag{2.2}$$

In this case P (Y | X) states the chance of data that has a vector X of a class Y, while P (Y) is an initial opportunity for class Y, whereas $\prod_{i=1}^{q} P(X_i|Y)$ is an independent class Y probability of all features in a vector X. The value of P (X) is always fixed so that in the predicted calculation we will only have to calculate the part P(Y), $\prod_{i=1}^{q} P(X_i|Y)$ by choosing the largest as the class chosen as a result of the prediction.

1.4. Stationary Time Series

The function of autocorrelation abbreviated as fak, formed by the set $\{\rho_k ; k = 1,2, ...\}$ with $\rho_0 = 1$. The autocorrelation in kth lag is defined as follows.

$$\rho_k = \frac{kov(Z_t, Z_{t-k})}{[\text{var}(Z_t).\text{var}(Z_{t-k})]^{\frac{1}{2}}} \quad (2.3)$$

An MA model is a model that observes and looks at the moving average of an event, then these events are used as observational data and are random. Generally the q level MA model is written in its mathematical form as follows.

$$Z_t = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + ... + \theta_q a_{t-q} \quad (2.4)$$

1.5. Artificial Neural Network

The arrangement model of a neuron that is included in the layer gives a good connection pattern that applies within and between layers often referred to as network architecture. Usually a single layer network consists of input neurons that are connected directly to a set of outputs on this network. Some models have only a neuron in the output layer.

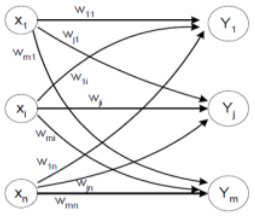

Figure 1. Single Layer Network

Figure 1 shows the network architecture with n input neurons $(x_1, x_2, \dots, x_n)$ and output neurons $(Y_1, Y_2, \dots, Y_m)$.

Activation function according to [13] is shown in Figure 1.

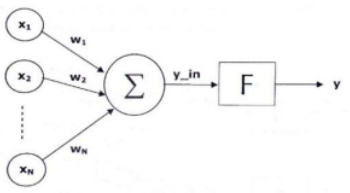

Figure 2. The activation function of a simple neural network

A neuron will process N inputs $(x_1, x_2, \dots, x_N)$ each of which has weights $w_1, w_2, \dots, w_N$ with the formula

$$y\_in = \sum_{i=1}^{N} x_i w_i \tag{2.5}$$

2. Methods

This research is a development in several researches and scientific works of the last few years. In a previous study [7] conducted a study in a doctoral dissertation scheme with the title "Fractional Integrated Recurrent Neural Network (FIRNN) for Time-Serving Data Forecasting with the Application of National Electricity Load". In this study the forecasting model was examined at a time using the RNN model. Whereas in 2015 and 2016 researchers got the opportunity in the grant scheme to compete with forecasting studies using recurrent neural network (RNN) on long memory and seasonal data, the forecasting results showed that the mean square error (MSE) using the RNN model was smaller than using the ARFIMA and ARIMA models. Developed a research on the productivity analysis of the performance of lecturers and educational staff in realizing the reputation of the Semarang State University (UNNES) using artificial neural networks [9].

The data used in this study are secondary data about the level of lecturer publications in the Faculty of MIPA UNNES, including scoupus indexed international journal publication data and have implications which are then referred to as X_1 variables, data of international journal publications are not scoupus indexed which are then referred to as X_2 variables, data on accredited national journal publications which are then referred to as X_3 variables, data on publications of non-accredited national journals which are referred to as X_4 variables, all of these data are obtained by using searches on SINTA and scoupus online.

Data analysis in this study uses several models and combines them to get better results and produce good statistics. Some models are used, among others, the existing data grouping is well presented using the Naïve Bayes model. Through the model, it is obtained a group of data that has good characteristics to be used as input data in pre-processing using a neural networking model. The results of processing data at the pre processing stage are then processed as training data, then the training data will be trained using several appropriate neural network models. After this process is completed, the final step is identification using a time series model to obtain the best identification results. The complete data analysis process can be described in the following flowchart.

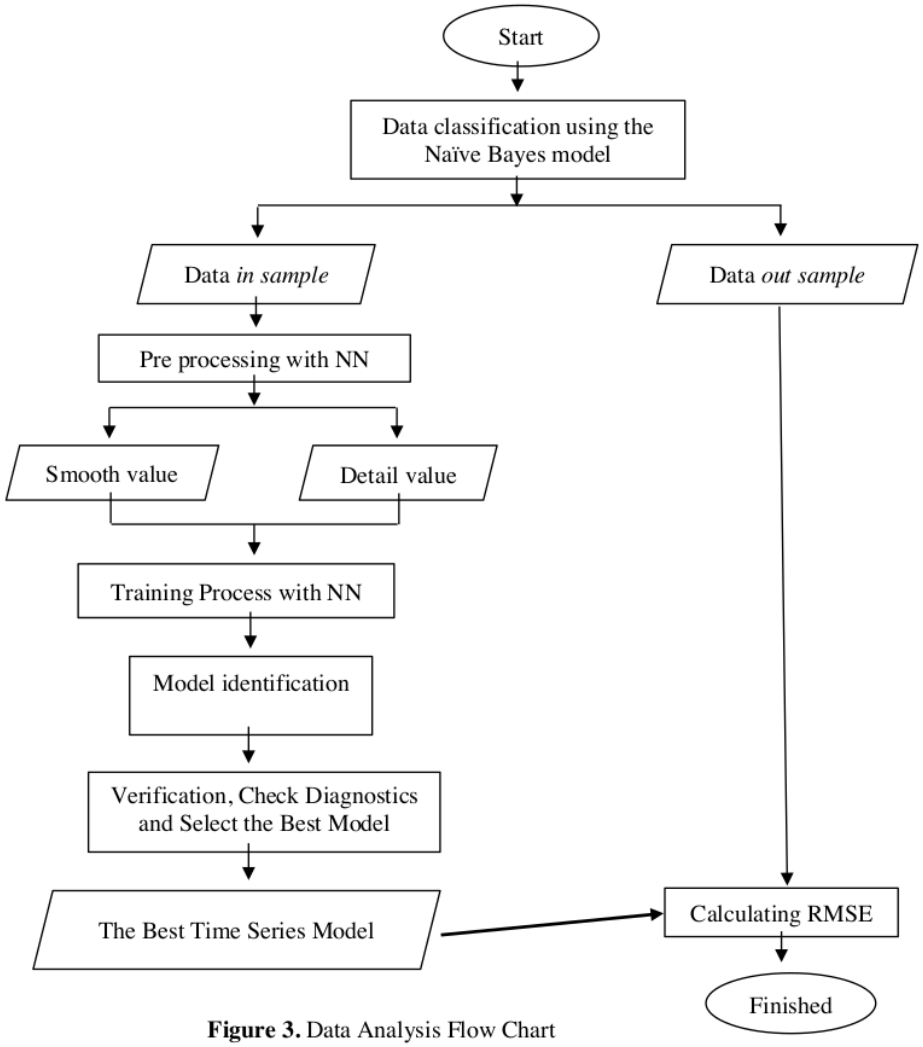

Figure 3. Data Analysis Flow Chart

3. Results and Discussion

Before pre-processing data with NN, first the data classification is done by using the Naïve Bayes algorithm, then preprocessing the data by normalizing the data using the following syntax.

[pn, ps] = mapstd(inputlatih) and
[tn, ts] = mapstd (target latih).

The command to build a BPNN network for data preprocessing, namely: net = newff (minmax (pn), [n 1], {'tansig' "purelin '},' traingdx ', while the newff command is used to build a BPNN network with n neurons in one hidden layer and one neuron at one output layer. The command " tansig " purelin '} shows the activation function used with tansig on hidden layers and purelin at the output layer. The traingdx command shows the training algorithm used in existing data, first Preprocessing is done by using transformation, in this study transformation is done on differing data in the range [1,1] and transformation of differential data in the range [-1,1].

The second transformation function in pre processing is by using prestd (inptrain) transformation with the aim of getting input data with normal distribution, namely achieving zero mean values and variance values is one. The results of the preprocessing stage using differing data through two models, namely the function model net = newff (minmax (in), [1,1], {'tansig', 'purelin'}) and the net = train function model (net, in, tn) The results obtained using the Minmax function model will then be used to conduct training for NN, in this study the BPNN model was taken.

Network training uses a momentum gradient descent training algorithm and an adaptive learning rate (trainminmax). The following will show the results of training on BPNN based on the best models of each momentum and learning rate (LR).

Table 1. Results of BPNN Training on data X1 with Momentum = 0.1 and LR = 0.01

Network Architecture	$f(x)$	*Epoch*	Perf	Training			Testing		
				MSE	MAPE (%)	R	MSE	MAPE (%)	R
2-25-1	minmax	1000	0,038	4,1452e+08	2,8196	0,877	3,3113e+09	8,6157	0,824

Based on Table 1. it can be seen that the best BPNN model on the gradient descent training algorithm with momentum and adaptive learning rate with momentum = 0.1 and LR = 0.01 was obtained on network architecture 2-25-1 in the logig activation function with MSE and MAPE the tests were 3.3113e + 09 and 8.6157% and the results were epoch 1000.

Table 2. Results of BPNN Training on data X2 with Momentum = 0.1 and LR = 0.01

Network Architecture	$f(x)$	*Epoch*	Perf	Training			Testing		
				MSE	MAPE (%)	R	MSE	MAPE (%)	R
2-25-1	minmax	1000	0,0292	4,1972e+08	2,9810	0,887	5,3712e+8	8,3527	0,82681

Based on Table 2, it can be seen that the best BPNN model in the gradient descent training algorithm with momentum and adaptive learning rate with momentum = 0.1 and LR = 0.02 was obtained on network architecture 2-25-1 in the logsig activation function with MSE and MAPE testing each of them is 5.3712e + 8 and 8.3527% and achievement is epoch 1000.

Based on Table 3. it can be seen that the best BPNN model on the gradient descent training algorithm with momentum and adaptive learning rate with momentum = 0.1 and LR = 0.03 obtained on network architecture 9-25-1 on the minmax activation function with MSE and MAPE the tests were 5.6229e + 10 and 8.5507% and the results were epoch 1000.

Table 3. Results of BPNN Training on data X_3 with Momentum = 0.1 and LR = 0.01

Network Architecture	$f(x)$	*Epoch*	Perf	Training			Testing		
				MSE	MAPE (%)	R	MSE	MAPE (%)	R
2-25-1	minmax	1000	0,0374	3,9427e+09	2,7371	0,9861	5,629e+10	8,5507	0,87891

Table 4. Results of BPNN Training on data X_4 with Momentum = 0.1 and LR = 0.01

Network Architecture	$f(x)$	*Epoch*	Perf	Training			Testing		
				MSE	MAPE (%)	R	MSE	MAPE (%)	R
2-25-1	minmax	1000	0,0329	5,1582e+4	3,5523	0,97817	4,0151e+8	9,1318	0,89664

Based on Table 4. it can be seen that the best BPNN model on the gradient descent training algorithm with momentum and adaptive learning rate with momentum = 0.1 and LR = 0.1 is obtained on network architecture 2-25-1 on the minmax activation function with MSE and MAPE the tests were 4,0151e + 8 and 9,1318% and the results were epoch 1000.

The results of the above training using BPNN show that the variables X_2 and X_4 provide relatively small MSE and MAPE values, this gives a perspective that scoupus international index publications and non-accredited national journal publications are the two most popular publications by FMIPA lecturers to realize publications in delivering internalization of the UNNES.

The next step by using the time series model will be model estimation and diagnosis. Based on ACF and PACF values, the possible ARIMA models are ARIMA (1, 2, 0), ARIMA (1, 2, 1) and ARIMA (0, 2, 1). The best model estimation results are ARIMA models (1, 2, 0). Based on the final parameter estimation obtained AR = -0.5804 and constant = -0.0198. So the model for ARIMA (1, 2, 0) is as follows

$$\phi_1(B)(1-B)^2 Z_t = \theta_0(B)a_t$$
$$(1-\phi_1 B)(1-2B+B^2)Z_t = a_t$$
$$(1-2B+B^2-\phi_1 B+\phi_1 B^2)Z_t = a_t$$
$$(1-(2+\phi_1)B+(1+\phi_1)B^2)Z_t = a_t$$
$$Z_t = (2+\phi_1)Z_{t-1} - (1+\phi_1)Z_{t-2} + a_t$$
$$Z_t = -0.0198 + (2-0.5804)Z_{t-1} - (1+0.5804)Z_{t-2} + a_t$$
$$Z_t = -0.0198 + (1.4196)Z_{t-1} - (0.4196)Z_{t-2} + a_t$$

The best results for identification related to the development of the type of scientific publications of lecturers is the best model is ARIMA (1, 2.0) because the parameter significance test, residual normality test, all independence tests and MSE values are small with the model that is $Z_t = -0.0198 + (1.4196)Z_{t-1} - (0.4196)Z_{t-2} + a_t$.

4. Conclusion

Based on the results of the research and discussion above can be concluded. The model of identification of lecturer publications in realizing internationalization of UNNES uses the Naïve Bayes classifier time series neural network algorithm through several stages, including data classification with the Naïve Bayes classifier algorithm model, then the pre-processing process using the neural network, the data training process using BPNN and ending with use a time series model to determine the best model obtained. The results of the identification of lecturers' publications using the Naïve Bayes classifier algorithm time series neural network resulted in non-indexed international journal publications and non-accredited national journal publications are the two most popular publications by FMIPA lecturers to realize publications in implementing internalization of UNNES.

References

[1] Haykin H 1999 Neural Networks: A Comprehensive Foundation, 2nd edition (rentice-Hall-Oxford)
[2] Menezes L M and McSharry P E 2006 *Int. J. Forecast.* **22** 1
[3] Ristiana Y 2008 Unpublished Final Project, Department of Physics Engineering, Institut Teknologi Sepuluh Nopember, Indonesia
[4] Swanson and White H 1995 *J Neural Comput.* **1** 425
[5] Hochreiter S, Bengio Y, Frasconi P and Schmidhuber J 2001 *J.F. (eds) A Filed Guide to Dynamical Recurrent Neural Networks*, IEEE Press
[6] Husen W 2001 Unpublished Final Project, Department of Physics Engineering, Institut Teknologi Sepuluh Nopember, Indonesia
[7] Walid, Subanar, Rosadi D and Suhartono 2015 *J Contemp. Eng. Sci.* **8** 1485
[8] Soni S 2011 *J InTech* **1** 10
[9] Walid and Alamsyah 2017 *J of Physics* **1** 25
[10] Olson and Delen 2008 *Advanced Data Mining Techniques* (SpringerVerlag- Berlin Heidelberg)
[11] Han J, Kamber M and Pei J 2012 *Data Mining: Concepts and Techniques (3nd ed)* (Waltham-Morgan Kaufmann)
[12] Lawrance and Balakrishma N 2001 *J of the Royal Statistical Society B* **63** 843.
[13] Kusumadewi S and Hartati S, 2010 *Neuro-Fuzzy: Integrasi Sistem Fuzzy & Jaringan Syaraf* (Graha Ilmu-Yogyakarta)

5

ORIGINALITY REPORT

12%	10%	6%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	pure.uva.nl Internet Source	3%
2	vital.lib.tsu.ru Internet Source	2%
3	Z.Q. Li, L. Ma, K. Khorasani. "A Dynamic Neural Network-based Reaction Wheel Fault Diagnosis for Satellites", The 2006 IEEE International Joint Conference on Neural Network Proceedings, 2006 Publication	1%
4	www.jbete.com Internet Source	1%
5	I H Diarsih, Tarno, A Rusgiyono. "Modeling of red onion production in Central Java using hybrid ARIMA-ANFIS", Journal of Physics: Conference Series, 2019 Publication	1%
6	lp3m.unuja.ac.id Internet Source	1%
7	www.warse.org	

Internet Source

1%

8 Nafizatus Salmi, Zuherman Rustam. "Naïve Bayes Classifier Models for Predicting the Colon Cancer", IOP Conference Series: Materials Science and Engineering, 2019
Publication

<1%

9 wiki.smu.edu.sg
Internet Source

<1%

10 Eko Nugroho Julianto, Ummu Salamah. "The students' interest for 2012 and 2013 cohort in construction engineering vocational education program Universitas Negeri Semarang in choosing the subject specialization", AIP Publishing, 2017
Publication

<1%

Exclude quotes	On	Exclude matches	< 10 words
Exclude bibliography	On		

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9