UMN Site Traffic Prediction using Technical Method

Alvin Alexander, Seng Hansun

Abstract*: The Marketing division of Universitas Multimedia Nusantara has a role in attracting prospective students to join and continue their next level of education. To help their task, the Marketing division needs an application that can predict the site traffic profile visits for the next following day. Therefore, a website-based application was created to assist the Marketing division. We used the Weighted Moving Average algorithm to calculate the number of subsequent visitors and its future prediction number. We also used the Weighted Mean Absolute Percentage Error after getting the prediction results to get the error rate of a prediction made. In the implementation process that has been done, the percentage error is 2.57%. The built application also had been evaluated by the UMN Marketing division employees by the means of questionnaires to find out how satisfied the application was.

Index Terms*: Marketing, Universitas Multimedia Nusantara, Weighted Moving Average, Weighted Mean Absolute Percentage Error.

I. INTRODUCTION

Universitas Multimedia Nusantara (UMN) is a university that has been established since 2006 with a definitive campus location in Kelapa Dua area of Summarecon Gading Serpong, Tangerang District, Indonesia. This university was inaugurated on November 20, 2006 [1].

Over time, the development of Universitas Multimedia Nusantara is very good from academic talks, profiles, history, and vision and mission. Therefore, Universitas Multimedia Nusantara also provides a website to access all of its contents from profiles, registrations, academics, student affairs, news, and achievements.

Prospective students from high school and vocational high school students in Indonesia would like to find information on universities throughout Indonesia. Every time a user visits the website, visitor data could be calculated based on Semrush. Semrush is a free website service that displays visitor statistics on a website, and there are also some paid services to get all of its access [2]. Semrush can search visitors based on referring page information, including search engines, advertisements, pay-for-click networks, email marketing, visiting country-based situations, as well as links contained in documents. All of these data are on the internet and can be accessed by creating a new account on the Semrush website.

Marketing is a social process and management where individual and groups obtain needs and desires through the process of creating and exchanging products and values [3]. Marketing division has the main role in promoting Universitas Multimedia Nusantara to get many prospective students to register at Universitas Multimedia Nusantara. Data owned by the Marketing division can be further used for visitor predictions, whether in days, weeks, or months count.

In the Moving Average algorithm, there are four main types, i.e., Simple Moving Average, Cumulative Moving Average, Weighted Moving Average, and Exponential Moving Average. Among the four types of algorithms, we chose the Weighted Moving Average because it is a part of the smoothing technique of the average group or moving average which has a weight factor calculation on it [4]. After that, it will be continued by evaluating the prediction results by using the Weighted Mean Absolute Percentage Error. To get the users (in this case, the Marketing division at UMN) satisfaction rates, we employ the End User Computing Satisfaction model as can be seen in the Results and Discussion section.

II. WEIGHTED MOVING AVERAGE

Forecasting is a very important element in making a decision. The effectiveness or failure of a decision generally depends on several factors, more precisely the critical success factors [5], that cannot be seen or known when the decision is taken.Preprocessing data can be done by the smoothing method. Smoothing is a process that is carried out on data so that it provides more precise interpolation data for subsequent calculations. Weighted Moving Average is a moving average method that has a weight factor calculation. But in WMA, there are weights used for each change in data, where the latest data has a greater weight value. Here is the formula for the Weighted Moving Average algorithm [4].

$$T_t = \sum_{-m}^{m} a_j y_{t+j} \quad (1)$$

$$a_j = \frac{Q(j.m)}{\sum_{-m}^{m} q(i,m)} \quad (2)$$

where

Y_{t+j} : the actual data
$Q(j,m)$: the weighted given to the data
T_t : *WMA* for the first period
$\sum q(i,m)$: summation of all weights

III. WEIGHTED MEAN ABSOLUTE PERCENTAGE ERROR

In testing the final predictions, we used the Weighted Mean Absolute Percentage Error (WMAPE). It is one measure of errors in forecasting cases. WMAPE is called the calculation of forecasting errors with weighted WMA,

Revised Manuscript Received on August 05, 2019

Alvin Alexander, Informatics Department, Universitas Multimedia Nusantara, Tangerang, Indonesia.
Seng Hansun, Informatics Department, Universitas Multimedia Nusantara, Tangerang, Indonesia.

where the value of this weight means the level of importance with other values, along with with with the formula used for the calculation of WMAPE [6] as shown in Eq. (3).

$$WMAPE(\%) = \frac{\sum\left|\frac{D_i - F_i}{D_i}\right| D_i}{\sum D_i} \quad (3)$$

where

D_i : Actual data to i

F_i : Prediction data to i

$D_i - F_i$: Error value in the period to i

Next, the calculation of the WMAPE error value will be used to calculate the performance using Eq. (4).

$$Performance = (1 - WMAPE)\, x\, 100\% \quad (4)$$

IV. END-USER COMPUTING SATISFACTION

After the implementation and development have successfully built, we use the evaluation technique to get user satisfaction by using the End-User Computing Satisfaction method (EUCS). EUCS has some variables, namely as follows: content, accuracy, format, ease of use, and timeless. Each element has questions that will be asked to the user. [7], [8]. The sample questions for this study can be seen in Table 1.

Table 1. Questionnaire Design

EUCS Dimension	Question ID	Question
Content	C1	The content of the information on the website predicts according to your needs?
	C2	The contents of the information on the prediction website are easily guaranteed?
	C3	The contents of the information on the prediction website are complete?
	C4	The contents of the information on the prediction website are clear?
Accuracy	A1	Have prediction website displayed correct and accurate information?
	A2	Every link in the prediction website that you click on always display the appropriate page?
Format	F1	Design of prediction website has attractive color settings?
	F2	Design of prediction website has a layout that makes it easy for users?
	F3	Design of prediction website has a menu structure and links that are easy to understand?
Ease of Use	E1	Is prediction website very easy to use?
	E2	Is a prediction website easily accessed from anywhere and at any time?
Timeliness	T1	Information about predictions that you need are quickly obtained through prediction website?
	T2	Does prediction website always display the latest information?

V. RESULTS AND DISCUSSION

A. Implementation Results

The following are a few pieces of screenshots of the application interface and their explanations.

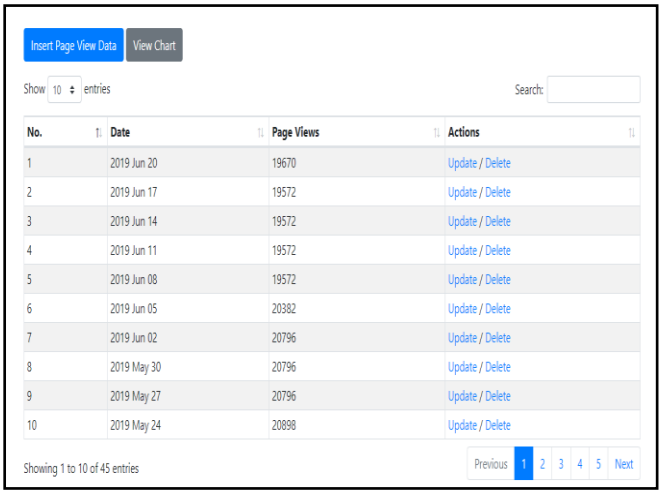

Figure 1. Home Page

The main page is shown in Figure 1. It is tasked with calling the data queries and equipped with features, such as "Insert Page View Data," "View Chart," "Update," and "Delete."

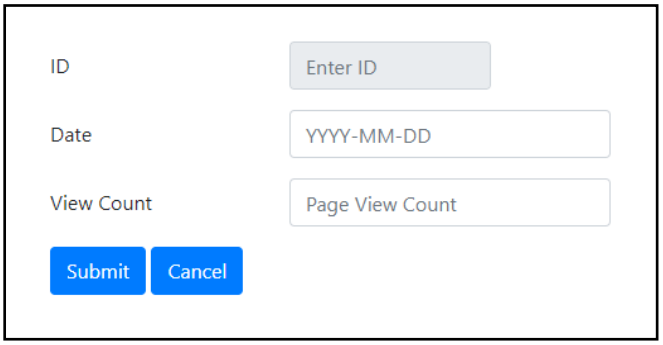

Figure 2. Form Page

The form page in Figure 2 task is to add data, change data, and delete data. The page will appear when the user selects "Insert Page View Data," "Update," and "Delete."

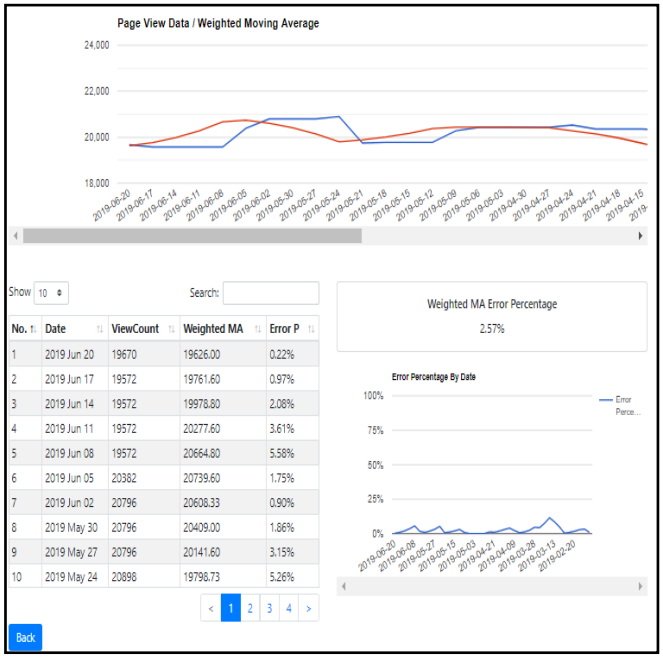

Figure 3. View Form and Results Page

Figure 3 explains the prediction results with the associated graph when a user selects the "View Chart" button. Two graphs will appear with different explanations, and a table appears to see all the predicted data.

B. WMA Simulation

The simulation takes five data on the first day. If you have got five days, then you can do the calculation. The first step is to determine the value based on a determined day, and each value each has a value of weights according to the period determined by the algorithm. The second step is the weight of the period determined based on the number of days predicted, the longer the day, the greater the weight value. Figure 4 explains the simulation of Weighted Moving Average calculation.

$$wma = \frac{((25938*1)+(24034*2)+(23567*3)+(23567*4)+(23567*5))}{1+2+3+4+5}$$

$$wma = \frac{356810}{15}$$

$$wma = 23787{,}33 \approx 23787$$

Figure 4. Weighted Moving Average Simulation

When calculating the data of the second position, the first day will be shifted to the second day and then the fifth day will be shifted to the sixth day, with no change in the weight of the rating. Figure 5 is the second calculation of the Weighted Moving Average simulation and will continue until the last day.

$$wma_2 = \frac{((24034*1)+(23567*2)+(23567*3)+(23567*4)+(23567*5)}{1+2+3+4+5}$$

$$wma_2 = \frac{353972}{15}$$

$$wma_2 = 23598{,}13 \approx 23598$$

Figure 5. Simulation of the Second Loop of Weighted Moving Average

C. WMAPE Simulation

In this simulation, the calculated WMA value will be calculated as a percentage error. The goal is to see how accurate the next day will be. The greater the percentage error, the more inaccurate the results are, and vice versa. Figure 6 describes the algorithm simulation when it has obtained WMA data, and Figure 7 describes the same simulation but with two WMA data on the next day.

$$wmape = \left(\frac{\left(\left|\frac{23567-23787}{23567}\right|\right)*23567}{23567}\right)*100\%$$

$$wmape = \left(\frac{220}{23567}\right)*100\%$$

$$wmape = 0.93\%$$

Figure 6. Weighted Mean Absolute Percentage Error Simulation

$$wmape = \left(\frac{\left(\left(\left|\frac{23567-23787}{23567}\right|\right)*23567\right)+\left(\left(\left|\frac{23567-23598}{23567}\right|\right)*23567\right)}{23567+23567}\right)*100\%$$

$$wmape = \left(\frac{220+31}{47134}\right)*100\%$$

$$wmape = 0.005\%$$

Figure 7. Simulation of Weighted Mean Absolute Percentage Error with two WMA data

D. EUCS Evaluation

This section will explain the evaluation results of End-User Computing Satisfaction. The application was demonstrated directly to UMN Marketing division employees located in Building A, 1st floor on June 21st and 22nd 2019. After some demo, we immediately gave the questionnaire form on June 22nd, 2019, and got four valid respondents along with the answers, as shown in Table 2.

Table 2. EUCS Questionnaire Response

Dimension	Q ID	SA	A	LA	DA	SDA
Content	C1	0	4	0	0	0
	C2	0	4	0	0	0
	C3	0	2	2	0	0
	C4	0	3	1	0	0
Accuracy	A1	0	3	1	0	0
	A2	0	2	2	0	0
Format	F1	1	2	1	0	0
	F2	0	2	2	0	0
	F3	0	3	1	0	0
Ease of use	E1	0	4	0	0	0
	E2	1	3	0	0	0
Timeliness	T1	0	3	1	0	0
	T2	1	2	1	0	0

Notes: SA (Strongly Agree), A (Agree), LA (Less Agree), DA (Disagree), SDA (Strongly Disagree)

VI. CONCLUSION

From the implementation results, it can be concluded that the Weighted Moving Average as a technical method in time series analysis has been successfully implemented in the form of a website application. Visitor traffic data based on Semrush internet source could produce WMA calculation and prediction results. Then WMAPE was used to calculate the percentage error with the value of 2.57%. The higher percentage error value inferred the more inaccurate prediction results for the next day. We also had evaluated the users' satisfaction rate by using the EUCS method. Overall, the users' satisfaction rate can be categorized as 'Agree' or acceptable, based on EUCS.

For future research, we could use other emerging technical methods for predicting website traffic. Some of the methods are Weighted Exponential Moving Average [9], Brown's WEMA [10], and Holt's WEMA [11].

REFERENCES

1. UMN, "Beranda." [Online]. Available: http://www.umn.ac.id/.
2. Semrush, "About Us." [Online]. Available: https://www.semrush.com/.
3. G. M. Armstrong, P. Kotler, M. J. Harker, and R. Brennan, Marketing: An Introduction. UK: Pearson, 2018.
4. F. Nhita, "IMPLEMENTASI ALGORITMA WEIGHTED MOVING AVERAGE PADA (FUZZY EAs) UNTUK PERAMALAN KALENDER MASA TANAM BERBASIS CURAH HUJAN," Indones. J. Comput., vol. 1, no. 1, Mar. 2016.
5. F. Costantino, G. Di Gravio, and F. Nonino, "Project selection in project portfolio management: An artificial neural network model based on critical success factors," Int. J. Proj. Manag., vol. 33, no. 8, pp. 1744–1754, Nov. 2015.
6. T. Wilson, "Short-term forecast error of Australian local government area population projections," Australas. J. Reg. Stud., vol. 21, no. 2, pp. 253–275, 2015.
7. R. Munap, S. N. B. Ahmad, S. A. Hamid, and M. F. B. M. Talib Beg, "The Influence of End User Computing System (EUCS) on User Satisfaction: The Case of a Logistic and Courier Service Company," Int. J. Soc. Sci. Humanit. Invent., vol. 5, no. 12, pp. 5103–5110, Dec. 2018.
8. E. Domínguez-Escrig, F. F. M. Broch, R. Lapiedra, and R. Chiva, "Promoting radical innovation through end-user computing satisfaction," Ind. Manag. Data Syst., vol. 118, no. 8, pp. 1629–1646, Sep. 2018.
9. S. Hansun, M. B. Kristanda, and P. M. Winarno, "Big 5 ASEAN capital markets forecasting using WEMA method," TELKOMNIKA (Telecommunication Comput. Electron. Control., vol. 17, no. 1, p. 314, Feb. 2019.
10. S. Hansun and S. Subanar, "Brown's Weighted Exponential Moving Average Implementation in Forex Forecasting," TELKOMNIKA (Telecommunication Comput. Electron. Control., vol. 15, no. 3, p. 1425, Sep. 2017.
11. S. Hansun and S. Subanar, "H-WEMA: A New Approach of DoubleExponential Smoothing Method," TELKOMNIKA (Telecommunication Comput. Electron. Control., vol. 14, no. 2, p. 772, Jun. 2016.

AUTHORS PROFILE

Alvin Alexander had just graduated from Universitas Multimedia Nusantara in 2019 and received his Bachelor degree in Informatics. He has participated in many events during his study at UMN and successfully finished it.

Seng Hansun had finished his Bachelor and Master degree from Universitas Gadjah Mada, majoring Mathematics and Computer Science program. Since 2011, he has been a lecturer and researcher at Universitas Multimedia Nusantara and published more than 75 papers both nationally and internationally. His research interests mainly in time series analysis and machine learning domain where he has successfully granted some research grants from the government and UMN institution.