Discrete Mean Amplitude of Glycemic Excursion (MAGE) Measurement on Diabetics with Spline Interpolation Method

Lailis Syafa'ah¹, Mauridhi Hery Purnomo², and Setio Basuki³

¹Electrical Engineering, University of Muhammadiyah Malang, Indonesia ²Electrical Engineering, Institut Teknologi Sepuluh Nopember, Indonesia ³Informatics Engineering, University of Muhammadiyah Malang, Indonesia ¹lailis@umm.ac.id, ²hery@ee.its.ac.id, ³setio_basuki@umm.ac.id

Abstract: Diabetes Mellitus (DM) is a disease that is characterized by glycemic disorders, including sustained chronic hyperglycemia and acute glucose fluctuations. Because DM is closely related to the body metabolism, the observation of the blood vessels becomes very important to perform. The observation is done by using the Mean Amplitude of Glycemic Excursion (MAGE). Definitively, MAGE is an important variable to solve clinical DM problems that contributes in generating oxidative stress related to the macro and microvascular complications. MAGE is technically used with continuous blood glucose data which is obtained by Continuous Glucose Monitoring (CGM). Because of the CGM is expensive for personal use, it cannot be used in the daily observation. The contribution of this study is the utilization of discrete data (3 days observation) to be used in MAGE measurement. This research employs Spline Interpolation technique to convert discrete blood glucose data to continuous signal. The validation of interpolated signal is conducted by comparing the pattern of discrete data and continuous signal for both original and clustered data. The experiment showed that both scenarios depicted identical pattern. The smallest RMSE was achieved by Linear Spline with 57.66 while the highest RMSE value was obtained by Quadratic Spline with 177.00.

Keywords: Continuous Glucose Monitoring, Diabetes Mellitus, Glycemic Disorders, Mean Amplitude of Glycemic Excursion, Spline Interpolation.

1. Introduction

Diabetes is a group of metabolic diseases that is characterized by hyperglycemia resulting from defects in insulin secretion or insulin action [1]. Theoretical and preclinical studies demonstrated that glycemic variability is an important parameter that is used to solve a clinical problem in patients with DM. Glycemic variability produces oxidative stress and potentially contributes to the development of macro and microvascular complications [2][3][4] especially heart diseases. Since the DM is closely related to the metabolism of the human body, it is paramount to monitor the blood vessels function. Mean Amplitude of Glucose Excursion (MAGE) as one of the glycemic variability is a method for measuring the fluctuations in blood glucose associated with body metabolism. In general, blood glucose fluctuations are observed continuously with Continuous Glucose Monitoring (CGM). As a readout device, the CGM provides mechanism to monitor blood glucose levels for every 5 minutes.

Delivering CGM for MAGE measurements for individual prediction purposes is difficult because the CGM is considered expensive. Therefore, this research proposes a model of blood glucose observation discretely. Since MAGE can only be measured with continuous signals [5][6][7], the observed discrete data must be converted into continuous form. The measurement in this research is conducted to monitor the discrete level of blood glucose 7 times a day during 3 days of observation. The contribution of this research is the preliminary study that discrete data is feasible to be used to measure MAGE. The experiment of this research is intended to obtain the most appropriate interpolation method of the blood glucose. The experimental results of this study will be useful for the development of MAGE-related studies in the future by using generic

Received: October 2nd, 2017. Accepted: May 30th, 2018 DOI: 10.15676/ijeei.2018.10.2.5 used blood-sampling devices. Interpolation is a way to find the value among several data points that have been known to estimate a function, in which that function is not defined by a formula but only by defined the data from the measurement results [8].

2. Mean Amplitude of Glucose Excursion (MAGE)

Several studies have shown the bad effects from sustained chronic hyperglycemia that results in excessive protein glycation and activates oxidative stress [2]. Role variability glucose (glycemic) is less documented, but the value of the average fluctuation of glucose will activate oxidative stress. Thus, it is recommended that the treatment strategies for diabetes should be directed to reduce to the minimum of the various components of Disglychemic. Glycemic variability is an important parameter that is used to resolve potential clinical problems in patients with diabetes. It is known that glycemic variability produces oxidative stress and potentially contribute to the development of macro and microvascular complications [3][4]. Currently, the best measurement for assessing glycemic variability is by Mean Amplitude of Glycemic Excursion (MAGE). However, MAGE not in routine clinical use. A routine measurement of glycemic variability clinical cause an important measure of overall glucose control. It predictors to the risk of complications of diabetes that are not detected by glycosylated hemoglobin levels (HbA1c). Good influence also reported by the DCCT (Diabetes Control Complications Trial) and by the UKPDS (United Kingdom Prospective Diabetes Study) 1998: Reduction of microvascular complications contribute to the reduction of glycemic excursion (glycemic variable). Assessment of glycemic control in diabetic patients should include three parameters, describes as "glucose triad" [3]: hemoglobin A1c (HbA1c), fasting plasma glucose (FPG), and postprandial glucose (PPG) as illustrated in Figure 1.



Figure 1. Three Glycaemia Parameters [9]

MAGE is a general size of the volatility of blood glucose levels, a indication level of diabetes control. Mage is usually used with a blood glucose monitoring system continuously (CGM). However, there were published medical studies which uses Mage algorithm by setting the data smaller, typically 7-10 observation per day for 2-3 days [4]. MAGE can be calculated by using equation 2.1 [10]. In this equation, the MAGE value is calculated with dividing λ (the series of blood glucose level) by *x* as number of observations.

$$MAGE = \sum_{r}^{\lambda}$$
(1)

There are two aspects that are important in the calculation of MAGE: first, not considering the MAGE excursion significant frequency, just average amplitude. Second, not considering the MAGE excursion glycemic which outside the normal range. MAGE only count high average amplitude of glucose that exceed the standard deviation for a particular day (only include the value of the peak to peak on bottom or vice versa) as shown in Figure 2 [4].



Figure 2. Example of MAGE calculation for 24 hours with CGM and SD=63

3. Materials and Method

This research is implemented based on several stages. The first step is to get blood glucose data from 60 volunteers. Patients who were willing to join the experiments will be divided into three groups: patients with normal, high, and very high blood glucose levels. The next step is to get the pattern of each blood glucose data for both discrete and continuous form. Therefore, it is necessary to implement the interpolation method using several variants of Spline technique namely Linear, Ouadratic, and Cubic. To strengthen the justification of using discrete to measure MAGE, the experiment is also done by clustering to see the centroid pattern on discrete data and continuous data. Details of the implementation and experiments are described in the sections below.

A. Patient Blood Glucose Observation

Table	1. Three Day	ucose	Sam	pling	Obse	rvati	on Ca	ırd	
				1.hour after Breakfast	Before Lunch	1 hour after Lunch	Before Dinner	1 hour after Dinner	Before Bedtime
Observation Time									
	Hyper Level	>300							
		260 - 299							
Blood Glucose Level		220 - 256							
		180 - 219							
		145 - 179							
		101 - 144							

Normal Level	80 - 100	 	 	 	
Hype Level	50 - 79	 	 	 	
	<50	 	 	 	

In this study, we conducted clinical experiments to take blood glucose data from patients (volunteers). The issue faced at this stage is how to identify patients who are classified as the normal, high, and very high-risk categories. Certainly, some patients are unaware of the risk of their blood glucose. Thus, patient selection is performed to ensure the distribution of blood glucose values through two types of activities. First, an interview was conducted to determine the patient's initial condition. The interview asks whether the patient has blood glucose-related problems. Second, each patient is taken blood samples at the clinical healthcare to take the patient's HbA1c (three months ago blood glucose level). If the distribution of HbA1c results is normal, then it can be continued to the next stage of taking blood glucose using a glucometer-accelerator for three days. If the HbA1c distribution is abnormal, it is necessary to obtain another patients will be taken blood samples that follow the seven-points observation scenario for three days.

Blood glucose sampling is conducted according certain scenarios: Before Breakfast, 1 hour after Breakfast, Before Lunch, 1 hour after Lunch, Before Dinner, 1 hour after Dinner, Before Bedtime. All of these scenarios are done for 3 days. Table 1 above shows the observation card to record the blood glucose value of each volunteer.

B. Interpolation Method

Basically, this research was to develop a system through software integrated with existing tools (glucometer). This system can calculate the value of glycemic variability. MAGE is the glycemic variability that can be calculated using this system. MAGE is usually show as continuously blood glucose monitoring system using a tool called CGM (Continuous Glucose Monitoring) as shown in Figure 3. However, due to technical reasons CGM cannot be done, then the measurement of glucose in this study conducted in discrete (non-continuous) to measure blood glucose as much as seven times a day for 3 days. Obtained data measurement is still in discrete data and must be converted into continuous data interpolation technique approach (ICGM / Interpolated Continue Glucose Monitoring), thus, the generated data model is same as CGM. Glucose levels on diabetic patients was measured using a glucometer Accu Check Performa Nano. It consists of a sensor (strip) that functions convert glucose into signals or electrical quantities and displayed on the monitor in the form of numerical data (discrete) that can be stored in a memory integrated with the tool. The data measurement is converted into data continue using spline.

As a standard for performing interpolation, this research uses 864 points that is taken from how the CGM obtain the continuous data. If in CGM sends data for every 5 minutes, then in an hour it will transmits 12 data. Then, there will be 288 data per day. Thus, within 3 days there will be 864 data.



Figure 3. Continuous Glucose Measurement [8]

The problem often faced in engineering is to estimate a value between several data points that have been known the value. The method used to estimate this is interpolation. We used 3 types spline interpolation method that are Linear Spline, Quadratic Spline, and Cubic Spline.

- Linear Spline

Linear Spline Interpolation (first-order Spline) is used to find the value of a point by connecting ordered points by using Linear Function [12]. This technique is the simplest piece of polynomial function. The resulting interpolation curve is relatively similar to the non-linear interpolation model. The curve contains a steep slope change in the data. The linear spline formula for the ordered data x_0 , x_1 , x_2 , ..., x_n is as follows:

$$f(x) = f(x_0) + m_0(x - x_0) \qquad x_0 \le x \le x_1$$

$$f(x) = f(x_1) + m_1(x - x_1) \qquad x_1 \le x \le x_2$$

...

$$f(x) = f(x_{n-1}) + m_{n-1}(x - x_{n-1}) \qquad x_{n-1} \le x \le x_n$$

(2)

with (slope) straight line *m_i*:

$$m_i = \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}$$

- Quadratic Spline

If the Linear Spline uses Linear Function formula to connect each point, then the Quadratic Spline Interpolation (second-order Spline) is used quadratic curve pattern. The quadratic spline formulas between points $(x_i, f(x_i))$ and $(x_{i+1}, f(x_{i+1}))$ are:

$$f_i(x) = a_i x^2 + b_i x + c_i$$

For n+1 points, there are *n* intervals, thus there exist 3n coefficient that are a_i, b_i, c_i i = 0, 1, 2, 3, ..., n [9]. Thus, there are 3n equations [13]:

Spline intersects data points by following this equation:

$$f_i(x_i) = a_i x_i^2 + bx_i + c_i = f(x_i) \qquad i = 1, 2, 3, \dots, n-1$$
(4)

$$f_i(x_{i+1}) = a_i x_{i+1}^{-1} + b x_{i+1} + c_i = f(x_{i+1}) \quad i = 1, 2, 3, ..., n-1$$
(5)
Spline continuous at the data point, expressed by the equation:

$$2a_{i}x_{i} + b_{i} = 2a_{i+1}x_{i+1} + b_{i+1} \qquad i = 1, 2, 3, \dots, n-1$$
(6)

The last condition formed freely as the second derivative of the spline between the first two data points to zero. Since the first instance for the first spline is $2a_1$ expressed

 $a_i = 0. (7)$

- Cubic Spline

Furthermore, cubic spline interpolation is a third-order (cubic) polynomial that works by deriving a third-order polynomial for each interval [12]. Formally, cubic spline interpolation is expressed by:

 $f_i(x) = a_i x_i^3 + b_i x^2 + c_i x + d_i = f(x_i) \qquad i = 0, 1, 2, ..., n$ (8)

Thus, for n + 1 data points, there are *n* intervals and 4*n* constants. As with the quadratic spline, the 4*n* condition requires the following evaluation [12]: (1) function value must be same at interior knots (2*n* – 2 condition), (2) The first and last functions must pass through the end points (2 conditions), (3) The first derivatives at the interior knots must be equal (*n* - 1 conditions), (4) The second derivatives at the interior knots must be equal (*n* - 1 conditions), and (5) The second derivatives at the end knots are zero (2 conditions).

C. Dynamic Time Warping (DTW)

To measure the distance between 2 time series curves x = [x1, x2, ..., xn] dan y = [y1, y2, ..., yn] can use Euclidean Distance method. The method is simply compute the distance by following the formula:

 $dist(x,y) = ||x-y|| = \sqrt{(x1-y1)^2 + (x2-y2)^2 + \dots + (xn-yn)^2}$ (9)

The problem arises in some conditions, suppose if x and y do not have the same length or fluctuations of the time series that because Euclidean Distance cannot find the best alignment. This problem can be solved with Dynamic Time Warping (DTW) try to find the minimum distance between vector of time series using Dynamic Programming (DP) [14]. There are two vectors t and r of length m and n. DTW is used to search for mapping path $\{(p1,q1), (p2,q2), ..., (pk,qk)\}$ with the aim of minimizing distance on $\sum_{i=1}^{k} |t(pi) - r(i)|$ with restriction: (1) (p1,q1) = (1,1), (pk,qk) = (m,n), (2) for every node (i,j) on path, there is a restriction that (i - 1, j), (i, j - 1), (i - 1, j - 1). Thus, to obtain the optimum distance value is done based on Forward DP:

- Optimum-Value function is defined as D(i,j) as DTW-distance between t(1:i) and r(1:j)
- Recursion Function:

$$D(i,j) = |t(i) - r(j)| + min \begin{cases} D(i-1,j) \\ D(i-1,-1) \\ D(i,j-1) \end{cases}$$

Initial condition D(1,1) = |t(1) - r(1)|

• Final answer: D(m, n)



According to [15], "The aim of the K-means algorithm is to divide M points in N dimensions into K clusters so that the within-cluster sum of squares is minimize". The algorithm is a clustering technique which follow certain steps starting with k groups each of which consists of a single random point, and thereafter adding each new point to the group whose mean the new point is nearest. After a point is added to a group, the mean of that group is adjusted in order to take account of the new point until convergence. Convergence is a condition where there is no change of centroid anymore. The distance that can be used is the Euclidean distance described in the previous section. The K-Means algorithm used in this research is used to obtain the centroid from the observation data of both observational discrete data and the interpolated data. Centroid formed from the two data will be compared to justify the similarity pattern.

4. Result and Discussion

A. Discrete Blood Sampling Observation

The results of blood glucose sampling should be recorded by the patients and the accompanying personnel. The Table 2 below represents discrete blood sampling observation results from 60 volunteers (patients). The main difficulties in this observation is that the patients forgot to record the blood glucose level according to the suggested time and the inconsistency of the patients to take the blood sampling. If these conditions appear, then the observation should be repeated on the next day. Therefore, the role of personnel who monitor this observation becomes critical.

Table 2. 3-Days Blood Glucose Sampling Observation									
Observation		Patient 1		Patient 2		Patient		Patient 60	
		Time in	Blood	Time in	Blood	Time in	Blood	Time in	Blood
		Minute	Glucose	Minute	Glucose	Minute	Glucose	Minute	Glucose
	0	0	100	0	100			0	200
	1	326	123	335	129			364	226
	2	398	217	448	198			595	298
	3	753	178	662	116			950	250
Day 1	4	858	227	783	236			1096	326
-	5	1105	137	1063	101			1187	304
	6	1191	149	1206	225			1320	310
	7	1368	161	1313	183			1543	232
	8	1805	98	1910	108			1893	243
	9	1878	309	2054	163			2047	256
	10	2184	115	2176	104			2302	192
Day 2	11	2262	113	2307	157			2429	257
	12	2558	78	2470	116			2590	249
	13	2819	152	2490	158			2724	290
	14	2826	128	2762	164			2946	217
Day 3	15	3213	118	3325	115			3225	244
	16	3319	245	3482	174			3359	301
	17	3596	113	3603	127			3837	213
	18	3660	142	3775	240			4000	284
	19	4037	116	3934	146			4082	275
	20	4177	128	4103	165			4220	210
	21	4369	128	4187	133			4330	219

T11 0 0 D D1 1 01

B. Blood Glucose Interpolation



Figure 5. (a) Measurement results of discrete data and (b) Measurement results of Interpolated Data

The discrete observed blood glucose information of one patient was interpolated using one of proposed smoothing technique. Basically, this comparison was conducted for all patients' data. However, in this paper we only present one of them since the aim of this stage is to get the general insight that both discrete and continuous data are identic. According to the mentioned interpolation scenario, it was used 864 interpolation points that follow the capacity of CGM to produce the continuous signal for 3 days of observation. The Figure 5 below shows the comparison chart of these two of data.

The x-axis, 1 to 21, represents the number of blood sampling. The y-axis represents the concentration of blood glucose of the patients. The value of blood glucose fluctuates and tend to rise after 1-hour food consuming (day-1: observation number 2, 4, 6; day-2: observation number 9, 11, 13; and day-3: observation number 16, 18, 20). In the Figure 5 (b), it is clearly seen that continuous blood glucose results provide information about amplitude, frequency, and wave pattern. This result is considered to be similar with the continuous signal produced by CGM. To compare the generated graphs of discrete and interpolation data, validation is also done by clustering results of all data. Comparisons are made to the centroid of the graphs using K-Means with DTW-based distance calculation. The argument of using this approach is to simulate the real situation that blood glucose sampling is collected by CGM. The Figure 6 presents that the centroid pattern generated by both data show the similar pattern.



(b) Figure 6. Similar Pattern of Discrete Data and Interpolated Centroid Data

To determine the most appropriate interpolation method to be used, in the experiment section, we formed three clusters of time series data from all patients. In the three clusters, we counted the centroids. Thus, the term "three" is not a three point of each patient but the number of centroids. To measure RMSE, we chosen randomly four validation points for each centroid. This means that we have 12 validation points (the discrete original data consists of 22 points for each patient that are obtained from 3 days observation). Therefore, the number of validation points used to calculate the RMSE are more than half. The error was calculated from the distance between some of the discrete original data that have been taken with the value of the interpolation results. The Root Mean Square Error (RMSE) was used to assess the performance of each interpolation technique.

Tuble 5: Interpolation Method Comparison and RMDE								
VALIDATION CENTROID POINTS		LINEAR SPLINE	QUADRATIC SPLINE	CUBIC SPLINE				
378.00	ter 1	449.66	582.3	493.73				
266.33	Clus	364.42	400.12	342				

Table 3. Interpolation Method Comparison and RMSE

394.83		348.08	405.28	380.92	
385.00		314.33	121.65	286.37	
208.21		260.39	457.75	279.12	
204.00	Cluster 3 Cluster 2	257.75	325.15	260.42	
249.86		218.18	165.96	224.71	
284.21		214.29	-4.94	183.01	
156.53		212.57	348.89	238.11	
147.42		194.48	254.28	201.89	
188.89		163.74	132.2	167.03	
183.74		162.08	17.54	141.94	
RMSE		57.66	177.00	70.70	

The smaller RMSE values indicate that the interpolation method used is better for predicting the validation points. The smallest RMSE was achieved with Linear Spline techniques that reached 57.66 while the highest RMSE value was obtained by using Quadratic Spline with 177. The RMSE value generated with Cubic Spline is 70.7. In this case, Linear Spline is considered the best technique to interpolate the discrete blood glucose data.

5. Conclusion

Basically, MAGE is calculated with continuous blood glucose data which is obtained by Continuous Glucose Monitoring (CGM). The CGM is expensive for personal and daily use. In this research, we propose the usage of discrete data (3 days observation) to be used in MAGE measurement. There are 60 volunteers in this research and the blood sampling is done by using standard glucometer. We provided sampling card to record the glucose level according to the observation schedule. The first experiment was done with comparing the discrete original data and interpolated data and the second experiment was intended to measure the similarity of centroid pattern of both data. In selecting the best Spline Interpolation method, we toke randomly some of the original dataset (3 centroids) and predicted values of these missing centroids. From three interpolation methods that is used in this experiment, Linear Spline shows the smallest RMSE value that is 57.66. The smallest RMSE indicates that the method is the most appropriate in predicting the interpolated points on blood glucose data.

6. References

- [1]. A. D. Association, "DEFINITION AND DESCRIPTION OF DIABETES OTHER CATEGORIES OF GLUCOSE," *Diabetes Care*, vol. VOLUME 33, 2010.
- [2]. C. C. LOUIS MONNIER, "Glycemic Variability: Should we and can we prevent it?," *Diabetes Care*, vol. 31 (Supple, 2008.
- [3]. F. Zaccardi *et al.*, "Group of Signs : A New Method to Evaluate Glycemic Variability," vol. 2, no. 6, pp. 1061–1065, 2008.
- [4]. C. R. Marling, J. H. Shubrook, S. J. Vernier, M. T. Wiley, and F. L. Schwartz, "Characterizing Blood Glucose Variability Using New Metrics with Continuous Glucose Monitoring Data," vol. 5, no. 4, pp. 871–878, 2011.
- [5]. G. Fritzsche, K.-D. Kohnert, P. Heinke, L. Vogt, and E. Salzsieder, "The Use of a Computer Program to Calculate the Mean Amplitude of Glycemic Excursions," *Diabetes Technol. Ther.*, vol. 13, no. 3, pp. 319–325, 2011
- [6]. G. Marics *et al.*, "Evaluation of an open access software for calculating glucose variability parameters of a continuous glucose monitoring system applied at pediatric intensive care unit," *Biomed. Eng. Online*, vol. 14, no. 1, pp. 1–13, 2015

- [7]. X. Yu *et al.*, "Calculating the Mean Amplitude of Glycemic Excursions from Continuous Glucose Data Using an Open-Code Programmable Algorithm Based on the Integer Nonlinear Method," vol. 2018, 2018.
- [8]. S. E. Pav, (*UCSD Math 174*, *Fall 2004*), Version 0., vol. 11. La Jolla, CA: Department of Mathematics, MC0112, University of California at San Diego, 2012.
- [9]. A. Tjokroprawiro, M. R. Indra, and D. Sargowo, "Expert System for Blood Glucose Fluctuations Measurement Based on MAGE (Mean Amplitude of Glycemic Excursion) and HbA1c On Diabetic Using K-NN (Nearest Neighbor)," vol. 4, pp. 135–141, 2014.
- [10]. P. Nelson, "Translating Glucose Variability Metrics into the Clinic via Continuous Glucose Monitoring : A Graphical User Interface ... Translating Glucose Variability Metrics into the Clinic via Continuous Glucose Monitoring : A Graphical User Interface for Diabetes," no. September 2011, 2017.
- [11]. M. Cescon, "Linear Modeling and Prediction in Diabetes Physiology Department of Automatic Control," no. June, 2011.
- [12]. R. P. Canale and S. C. Chapra, *Numerical Methods for Engineers*, Sixth Edit. McGraw-Hill Higher Education, 2010.
- [13]. A. V. Ayanti, "Perbandingan Interpolasi dalam Metode SPLINE," Universitas Sanata Dharma Yogyakarta, 2007.
- [14]. R. Jang, "Data Clustering and Pattern Recognition," 2005. [Online]. Available: http://mirlab.org/jang/books/dcpr/. [Accessed: 28-Sep-2017].
- [15]. J. A. HARTIGAN, M. A. WONG, and Yale, "Algorithm AS 136 A K-Means Clustering Algorithm," J. R. Stat. Soc. Ser. C (Applied Stat., vol. 28, no. 1, pp. 100–108, 1979.



Lailis Syafa'ah (lailis@umm.ac.id) works at University of Muhammadiyah Malang (UMM) as a lecturer at Electrical Engineering Department. She received her Bachelor and Master Degrees from Electrical Engineering Department, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia, in 1989 and 2002 respectively. In 2014, she received Doctoral Degree form Medical Department, Universitas Brawijaya, Malang, Indonesia. Currently, she actively conducts research on Biomedical Engineering.



Mauridhi Hery Purnomo (hery@ee.its.ac.id) currently is a Professor at Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember Surabaya, Indonesia. His research interests are in Artificial Intelligent Technique for Power System and Biomedical Engineering.



Setio Basuki (setio basuki@umm.ac.id) was born in Indonesia, in 1984. He graduated from Institut Teknologi Telkom (ITTelkom), Indonesia, 2007, for his Bachelor Degree in Informatics Engineering. He received Master Degree in Informatics Engineering from Institut Teknologi Bandung (ITB), Indonesia, 2015. His research interests are the applications of Machine Learning and Natural Language Processing. Since 2009, he has worked as a lecturer in Informatics Engineering Department, University of Muhammadiyah Malang (UMM), Indonesia.