Machine learning-based characteristic identification of MSG content in gravy foods

Phisca Aditya Rosyady[1*], *Nurina Umy* Habibah[2], *Masita*[1], *Anton* Yudhana[1]

[1]Department of Electrical Engineering, Universitas Ahmad Dahlan, Bantul, Daerah Istimewa Yogyakarta, Indonesia
[2]Department of Health Nutrition, Universitas Gadjah Mada, Yogyakarta, Daerah Istimewa Yogyakarta, Indonesia

Abstract. Monosodium Glutamate (MSG) is a sodium salt that binds to amino acids in the form of glutamic acid, widely used as an additive in cooking as a flavoring. Therefore, this research aims to detect the level of MSG content in soupy foods using Machine Learning. This research determines the identification of MSG using the Machine Learning method Naive Bayes classifier algorithm in Python software. This tool determines the identification of MSG dissolved in water using a Photodioda sensor, push button, RGB LED, Arduino Nano and Resistor. From the research obtained the results that the color of the light source affects the sensor reading value. Sensor value readings based on different light sources have the same pattern, but different values. The difference in sensor value is caused by the effect of LED color on specimen color. The more MSG used, the greater the photodiode sensor reading value. Based on this research, the accuracy value is 83.6%.

1 Introduction

In the development of science and technology with an impact in various changes in lifestyle and community needs. One of them is the change in consumption patterns in the community tends to want delicious food in a practical way. Therefore, for the use of food additives that are very widely used L-acid glutamate compounds in the form of its salt, namely Monosodium Glutamate (MSG). On average, many people in Indonesia consume around 30 grams of MSG [1]. Monosodium Glutamate (MSG) is widely used to add flavor to food, consuming MSG in the long term can have adverse effects on the health of the human body such as damaging the brain, damaging reproductive function, liver, kidneys [2].

According to this study Monosodium Glutamate (MSG) has many characteristics when using fast food ingredients not from natural ingredients, it will be fatal to the health of the body, especially the nerves of the brain. If most use MSG can be dangerous. The content of glutamic acid has the highest portion in food, so in the process of cooking food there is no need to add MSG cooking flavoring. Many housewives state that if the dish does not use artificial flavoring or MSG, the taste of the dish will be less delicious and are accustomed to

* Corresponding author: phisca.aditya@te.uad.ac.id

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

adding MSG to the dish [3]. So, in order to avoid problems with health it is better to reduce the use of MSG [4]. Naive Bayes is a method that can be used to predict based on certain samples and is efficient in the data classification process, so as to reduce errors in data collection [5]. Bayesian is used to generate data retrieval estimates on soupy dishes by combining estimates from samples and other existing information. This data retrieval method has a learning process based on training data, using conditional probability as its basis [6]. The success rate of this method is highly dependent on the initial knowledge provided. The naive bayes classifier (NBC) method is a method used for text classification. Where NBC uses probability theory as a basic part of the theory. In his book, Han, J. And Kamber, M. Which states: "Bayesian classifiers have a high level of speed and accuracy when applied in large databases" (2001, p.296) [7]. Based on the background previously described, this research will build a tool to detect MSG in soupy foods that can be known using a photodiode sensor with analysis reinforced by the naive bayes classifier algorithm through the MSG recognition stage so as to get MSG characteristics that can be identified from soupy foods as a reference for applying the characteristics of the MSG [8].

In this study, an MSG detector was made using a Photodiode Sensor, especially for soupy foods that can be monitored remotely based on the Internet of Things (IoT) which is displayed on the Blynk software. In this test, the MSG solution is heated and mixed with Ninhydrin liquid reaction material to bind the amino acids contained in the MSG compound. then the voltage results are measured with the detector periodically within 30 seconds. The results of this study showed a decrease in the value of light emission (LUX) received by the Photodiode Sensor with an average decrease in value for spinach vegetable soup samples of 1713.07 LUX, and meatball soup of 2226.24 every addition of 0.2 grams of MSG [9].

Food production is one of the largest contributors to climate change. The contribution of the food industry to the SDGs has been of particular interest since the inception of the targets as food production is closely linked to many of the SDGs. Policies relating to the food industry have been developed to achieve maximum effectiveness, meaning that they aim to contribute to as many SDGs as possible with an overarching strategy for food production.

2 Material and Methods

2.1 Object

The object of research in this study is MSG. This research identifies MSG when MSG becomes a substance dissolved in pure water. MSG will be detected based on its turbidity level through a photodiode sensor. The level of turbidity will be used as a class, this research will be divided into four classes, namely pure, low, medium and high. This research uses the naive bayes classifier (NBC) method which has a high level of speed and accuracy when applied in large databases.

2.1 Block Diagram

The hardware design is made in a block diagram. The block diagram can be seen in Figure 1.

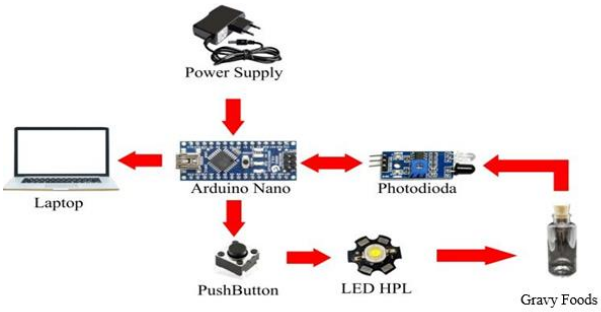

Fig. 1. System Block Diagram

The block diagram consists of 2 processes, namely input and output, from the push button which functions as input. LED serves as a light source while the Photodiode sensor as output. The photodiode sensor functions to detect the turbidity value of the MSG specimen used.

2.2 Flow Chart

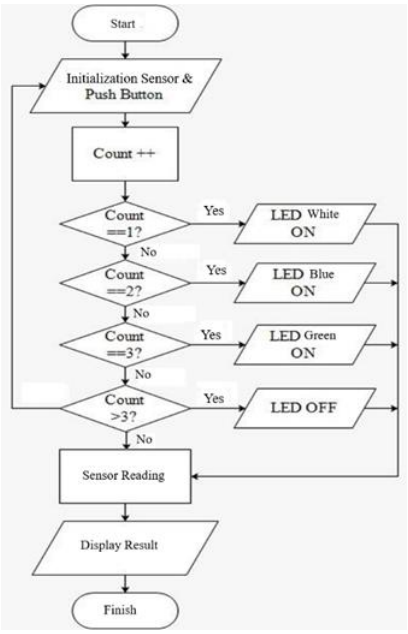

Fig. 2. Flow chart

In Figure 2, the detection system in MSG includes several inputs, processes and outputs. The input is a push button and photodiode sensor, the push button is used to control the color output of the LED and the photodiode sensor is used to absorb light from the LED. For the process part, when the push button is pressed or worth 1 then the LED will light up. At the process stage a count++ or calculation occurs. So, when count=1, the white LED is ON while the blue LED and green LED are OFF. When count==2, the blue LED is ON, while the white LED and green LED are OFF. When count==3, the green LED is ON while the white and blue LEDs are OFF. When count> 3, the three LEDs will turn off and count = 0. To display the data taken from the photodiode sensor, it will appear on the serial monitor on the laptop.

The output stage displays the value of the sensor readings and will perform data processing using the Machine Learning Algorithm Naive Bayes Classifier method. The flowchart of data processing in Naive Bayes Classifier is shown in Figure 3.

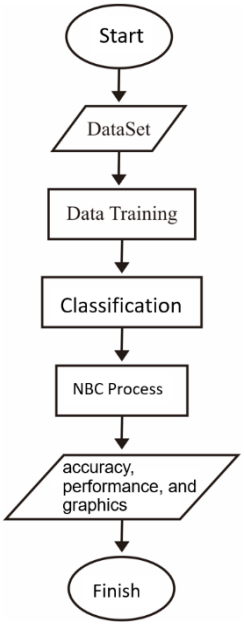

Fig. 3. *Naive Bayes Classifier* Flowchart

In Figure 3 the dataset acts as input in the Naive Bayes Classifier process. Then the data is processed with python by dividing it into training data and testing data. All data has been labeled with four classes namely pure, low, medium and high. The test data will be tested whether the identification results match the training data that has been created before. The output is accuracy, performance and graph.

3 Results and Discussion

3.1 Sensor Accuracy Testing

In this test we have tested the accuracy level of the photodiode sensor comparing the value generated on the photodiode used in this system with the value generated on the lux meter as a calibrated and credible measuring instrument.

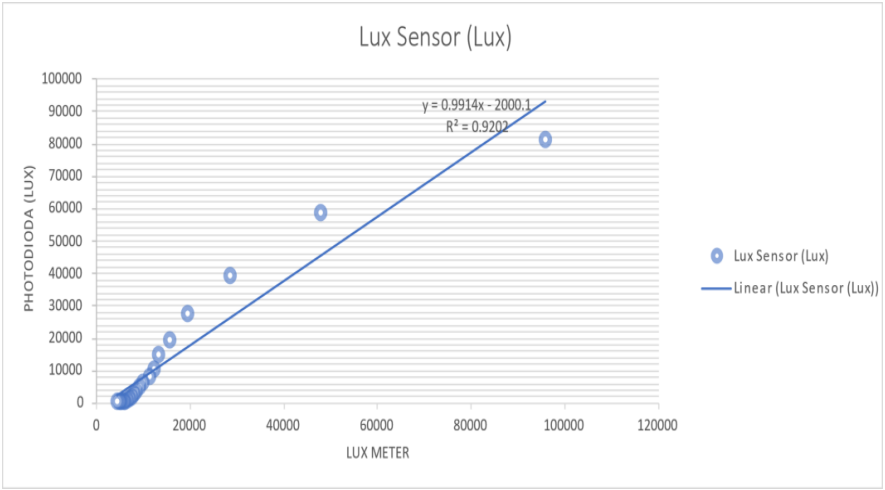

Fig. 4. Calibration Graph

Figure 4 shows a graph of the relationship between the lux meter reading value and the lux sensor. Obtained a graph covering the relationship with the line equation, namely y = 0.9914x - 2000.1 is a power regression model between the two variables with a value of R2 = 0.9202 referred to as the coefficient of determination which means the lux meter variable by the lux sensor variable is 92.02%. The remaining 7.98% is an error, then the remaining concentration is influenced by other variables.

3.2 Comparison of Photodiode Sensor Voltage Values Based on Different LEDs

This test is to get the photodiode sensor value of each LED color. The LEDs used are blue, green and white. MSG specimens are made from Ajinomoto brand MSG. The sample used as MSG specimen started from 100ml of pure water then added 0.5 gram for each test. Wates was heated to boiling for two minutes or twenty seconds. After boiling, the water is mixed with 1 ml of liquid ninhydrin and 0.1 gram of sodium acetate at a temperature of 70o to 80o. Ninhydrin is useful for changing the color of a liquid solute given amino acids into a purple color. There are several samples used and 30 data are obtained in each sample with three LED indicators used. So that the total data obtained is 320 data. The sensor value is obtained in real-time in the form of voltage. The maximum voltage value of the photodiode sensor is 5V, that means the sensor reading value of the photodiode will not even be more than 5V even if the sensor is not able to absorb light. From this study, the results obtained that the color of the light source affects the performance of the photodiode sensor. The graph of the test results is shown in Figure 5.

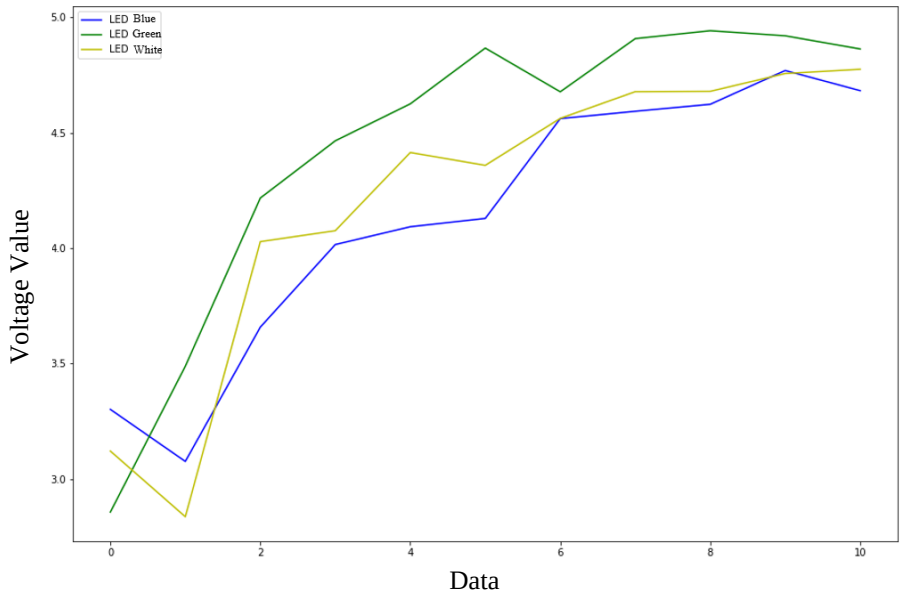

Fig. 5. Photodiode Test

Based on the voltage values shown in Figure 5, the blue line represents the white LED, the orange line represents the green LED and the green line represents the blue LED. From Figure 5, it is known that the value of the photodiode sensor with the green LED as the light source has a higher voltage value than the other LED values. The values generated by the photodiode sensor readings with all the LEDs as light sources are fairly constant it can be observed that the LEDs of almost all the LEDs have increased in each experiment. However, although they have almost the same pattern, the values obtained are different, this is also influenced by the mixing of color LEDs as light sources which are given a purple color produced by the chemical compounds in dissolved MSG. That way it is able to produce the presence or absence of spectral colors (other color combinations). In a previous study, it was stated that the difference in output voltage value is caused by the difference in wavelength of each LED. The value in the form of voltage value as a dataset is processed in Machine Learning using the Naive Bayes Classifier Algorithm.

3.3 Data processing using Naive Bayes Classifier (NBC)

This naive bayes classifier (NBC) classifier process can predict the probability of class membership of data that will fall into a particular class, according to probability calculations. naive bayes classifier (NBC) shows a high level of accuracy with simple calculations. The result of characterizing the number of input variables in the NBC process is several variables. NBC will transform the data into a new coordinator. From these variables will be reduced or simplified into several new coordinate dimensions, namely precision, recall, f1-score and support coordinates. The formation of new coordinates in the form of precision, recall, f1-score and support will produce a significant value variance from the total data variance and is used to create visualization charts in several dimensions. In the NBC process used to create the results of the difference in patterns from the four MSG categories. However, based on the accuracy produced, which is quite large at 83.6%, this research can be said to have achieved its goals. To find out the performance of the tests we have done, we can see the results of data analysis using the NBC algorithm in Figure 6.

	precision	recall	f1-score	support
0	1.00	1.00	1.00	15
1	0.92	0.50	0.65	24
2	1.00	0.95	0.97	19
3	0.54	1.00	0.70	14
4	1.00	1.00	1.00	16
5	0.90	0.35	0.50	26
6	1.00	0.94	0.97	17
7	0.46	1.00	0.63	11
8	0.82	1.00	0.90	18
9	1.00	1.00	1.00	13
10	1.00	1.00	1.00	16
accuracy			0.84	189
macro avg	0.88	0.88	0.85	189
weighted avg	0.89	0.84	0.83	189

Fig. 6. Algorithm Performance

In Figure 6 the first column contains labels from the state or identification level 0 indicates pure water, 1 to 10 indicates low, medium, and high levels. The figure shows the relationship between the values of precision, recall, f1-score and support for each label. The next column has precision, precision is the similarity between the test data and the information needed. Recall is the level of success of the program in finding information. Accuracy is the level of truth or an approach from the tested value to the real value that has been displayed in Figure 7.

```
[[15  0  0  0  0  0  0  0  0  0  0]
 [ 0 12  0 12  0  0  0  0  0  0  0]
 [ 0  1 18  0  0  0  0  0  0  0  0]
 [ 0  0  0 14  0  0  0  0  0  0  0]
 [ 0  0  0  0 16  0  0  0  0  0  0]
 [ 0  0  0  0  0  9  0 13  4  0  0]
 [ 0  0  0  0  0  1 16  0  0  0  0]
 [ 0  0  0  0  0  0  0 11  0  0  0]
 [ 0  0  0  0  0  0  0  0 18  0  0]
 [ 0  0  0  0  0  0  0  0  0 13  0]
 [ 0  0  0  0  0  0  0  0  0  0 16]]
Accuracy of the model:83.6%.
```

Fig. 7. Testing Accuracy

Based on Figure 7, the accuracy value is 83.6% of the total data variation, which carries all the information contained in the data. based on the results of the test accuracy value can form a graph in Figure 8.

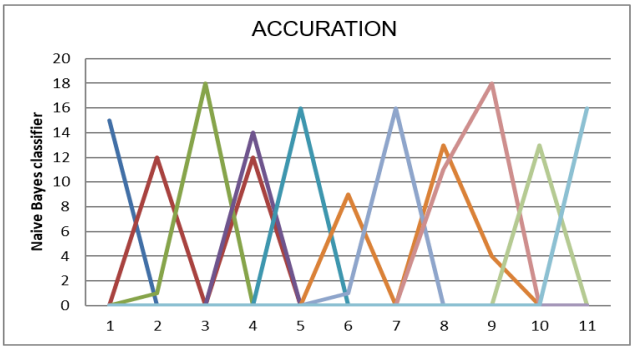

Fig. 8. Graph of Naive Bayes classifier

Figure 8 shows the results of testing accuracy using the Naive Bayes classifier algorithm, that grouping can be done based on the category used from the accuracy value.

4 Conclusion

After designing, making and testing the research on identifying the characteristics of MSG content in soupy food based on Machine Learning, which uses MSG specimens starting from pure ari and there is an addition of 0.5 grams in each test up to 5 grams. Then the results of testing the accuracy of the sensor by calibrating can be obtained the value of the graph equation, namely y = 0.9914x - 2000.1 is a power regression model between the two variables with a value of R2 = 0.9202 referred to as the coefficient of determination which means the lux meter variable by the lux sensor variable is 92.02% the remaining 7.98% is an error. For the results of the analysis using the Naive Bayes Classifier Algorithm produces an accuracy value of 83.6%.

References

1. I.N. Adiyasa, L.K. Abdi, & R. Fujiawati, Tingkat pengetahuan ibu, peran petugas kesehatan dan perilaku penggunaan penyedap rasa Monosodium Glutamat (MSG) pada masakan. J. Kesehat. Prima **10**, 1756 (2016)
2. T. Ayudia, Sistem pakar diagnosa penyakit akibat konsumsi berlebihan Monosodium Glutamat (Msg) menggunakan metode anfis. J. Pelita Informat. **8**, 382, (2020)
3. M. Bari, S.H. Sitorus, & U. Ristian, Implementasi metode naïve bayes pada aplikasi prediksi penyebaran wabah penyakit ISPA (Studi Kasus: Wilayah Kota Pontianak), 06(03), 205 (2018)
4. A.N. Respati, Pengaruh penggunaan pasta labu kuning (*Cucurbita Moschata*) untuk substitusi tepung terigu dengan penambahan tepung Angkak dalam pembuatan mie kering. Skripsi, Universitas Sebelas Maret, Surakata, (2010)
5. N. Cebi, T. Ozturk, C.E. Dogan, & O. Sagdic, Evaluation of raman microscopy for the detection of additional monosodium glutamate in dry soup mix. Qual. Assur. and Saf. of Crop. & Foods **12**, 1 (2020)
6. P.I. Cahyani, Hubungan konsumsi makanan mengandung Monosodium Glutamate (Msg) dengan perkembangan kognitif anak usia sekolah di SD Negeri Pongtiku 2 Kota Makassar. Sekolah Tinggi Ilmu Kesehatan Panakkukang Makassar, (2020).

7. A.M. Chikmah, & J. Nisa, Pengaruh lifestyle (pola konsumsi makanan mengandung Msg) terhadap gangguan pemusatan perhatian dan hiperaktif pada anak prasekolah. BHAMADA: J. Ilmu. Teknol. Kesehat. **11**, 65 (2020).
8. D.R. Adawiyah, & F. Setiawan. Ambang deteksi dan preferensi rasa umami dalam model pangan. J. Teknol. dan Indus. Pang. **28**, (2017)
9. F. Liantoni, & A.A. Santoso, Penerapan ciri statistik orde pertama dengan Ekualisasi histogram pada klasifikasi telur omega-3**.** Simetris: J. Tek. Mesin, Elektro. dan Ilm. Komp. **9**, 953 (2018).
10. F.W. Ningtyias, A.H. Asdie, M. Julia, & Y.S. Prabandari, Makanan mentah, goitrogenik dan gangguan akibat kekurangan yodium (Gaky)(Raw Food, Goitrogenic and Idd). Bul. Penelit. Sist. Kesehat. **18**, 20933 (2015).
11. M. Lee, J.W. Jung, D. Kim, Y.J. Ahn, S. Hong, & H.W. Kwon, Discrimination of umami tastants using floating electrode-based bioelectronic tongue mimicking insect taste systems. Acs Nano **9**, 11728 (2015).
12. J. Liu, Y. Fan, G. Chen, & Y. Liu, Highly sensitive glutamate biosensor based on platinum nanoparticles decorated MXene-Ti3C2Tx for L-glutamate determination in foodstuffs. Lwt **148**, 111748 (2021)
13. Y. Mirza, Sensor suhu Lm35 dan photo dioda sebagai sistem kendali mesin potong. JUPITER J. Penel. Ilm. dan Teknol. Komp. **10**, 45 (2018)
14. A. Mulyasuryani, & D.E.D Prasetya, Development of chemical sensor for detection of monosodium glutamate by polyvinyl alcohol-Fe3O4 membrane on screen printed carbon electrode. In *IOP Conference Series: Mater. Sci. and Engineer.* **546**, 032022 (2019)
15. W. Muslehatin, & M. Ibnu, Penerapan naïve bayes classification untuk klasifikasi tingkat kemungkinan obesitas mahasiswa sistem informasi UIN Suska Riau, In *Seminar Nas. Teknol. Inform. Kom. dan Indus.* 18 (2017)
16. M.T. Rizki, Pemantauan status kualitas air berbasis Extensible Messaging and presence protocol (XMPP) dengan metode klasifikasi Naive bayes. (2018)
17. E. Murdiana, Analisis penggunaan monosodium glutamat (MSG) pada ibu rumah tangga di perkotaan dan perdesaan Bogor. Institut Pertanian Bogor, Bogor. (2012)
18. O.D. Nurhayati, D. Eridani, & A. Ulinuha, Ekstraksi ciri orde pertama dan metode principal component analysis untuk mengidentifikasi jenis telur ayam kampung dan ayam arab. J. Sist. Infor. Bisn. **9**, 133 (2019)
19. M.K. Rahmadhika, & A.M. Thantawi, Rancang bangun aplikasi face recognition pada pendekatan CRM menggunakan opencv dan algoritma haarcascade. IKRA-ITH INFORMATIKA: J. Komp. dan Inform. **5**, 109 (2021)
20. A. Syariffudin, & U. Pratiwi, Perancangan alat peraga gerak harmonik berupa bandul matematis menggunakan sensor photodioda berbasis arduino. J. Pendidik. Sains dan Komp. **2**, 2809 (2022)
21. M. Sabransyah, Y.N. Nasution, & D. Tisna, Aplikasi metode naive bayes dalam prediksi risiko penyakit jantung naive bayes method for a heart risk disease prediction application. J. EKSPONENSIAL **8**, 111 (2017)
22. D. Sartika, & D. Sensuse, Perbandingan algoritma klasifikasi naive bayes, nearest neighbour, dan decision tree pada studi kasus pengambilan keputusan pemilihan pola pakaian, JATISI J. **3**, 151 (2017)
23. Surantoro, et al. Pembuatan alat percobaan hukum II newton dengan sensor cahaya

photodioda. J. Mater. dan Pemb. Fisik. **12**, 1-5 (2020)
24. UCI Machine Learning Repository, 2018, Center for Machine Learning and Intelligent Systems, UCI Machine Learning Repository, https://archive.ics.uci.edu/ml/datasets.htm, Diakses Jpada tanggal 20 Juni 2018.
25. P.D Wananda, L. Novamizanti, & R.D. Atmaja, Sistem deteksi cacat kayu dengan metode deteksi tepi SUSAN dan ekstraksi ciri statistik. ELKOMIKA: J. Tek. Ener. Elek. Tek. Telekom. & Tek. Elek. **6**, 140 (2018)
26. Wulandari. Senyawa flavor hasil hidrolisis dari ikan seluang (*Rasbora* sp) sebagai sumber MSG (*Monosodium glutamate*) alami. In *Prosiding Seminar Nasional Lahan Suboptimal* 2019, Palembang 4-5 September 2019. **496**. Palembang: Unsri Press. (2019)