Ranking Universities in Medan Using WoE and IV in Weighting of MAUT

Putri Maulidina Fadilah[1]*, Adidtya Perdana[2], Nurul Ain Farhana[3]
[1][2][3]Universitas Negeri Medan, Indonesia
[1][3]Department of Statistics [2]Department of Computer Science
[1]putrimaulidina@unimed.ac.id, [2]adidtya@unimed.ac.id, [3]nurulainfarhana@unimed.ac.id

Abstract: Throughout Indonesia, including Medan, the popularity of a university can be specified by the ranking of a university. There are five assessment components which important for ranking universities under the Ministry of Education, Culture, Research, Technology and Higher Education, such as the Quality of Human Resources, Institutional Quality, Quality of Student Activities, Quality of Research and Community Service, and Quality of Innovation. Multi Attribute Utility Theory (MAUT) is one of the decision support system (DSS) methods that can be used to calculate campus rankings. However, the researcher were determining the weight of MAUT method based on their preferences and it was subjective. Weight of Evidence (WoE) can be used to assign a numerical score to each category of independent variables that describes the strength of its relationship to the target variable. In selecting the independent variable that is most informative and relevant in predicting the target variable, Information Value (IV) can be used. Based on the results, college B is the most popular university out of ten universities in Medan, with the highest evaluation value 0.796296296 using MAUT method and 0.923794719 using MAUT method with WoE & IV. The last position is J college with the lowest evaluation value 0.1666666667 for MAUT method and 0.02540176 for MAUT method with WoE & IV. The weighting of MAUT method with WoE and IV produces more optimal evaluation value than the the original MAUT method.

Keywords: Decision support system; information value; multi attribute utility theory; ranking university; weight of evidence

INTRODUCTION

The ranking of a university, can be used by highschool students as a reference for choosing the university. The popularity of a college can be specified by the ranking of a university. The clustering scheme for higher education under the Ministry of Education, Culture, Research, Technology is needed to determine the ranking of a university. The ranking can also enhance universities to improve their quality of higher education continuously.

From 2018, there are five assessment components which important for ranking universities under the Ministry of Education, Culture, Research, Technology such as the Quality of Human Resources, Institutional Quality, Quality of Student Activities, Quality of Research and Community Service, and Quality of Innovation. Human Resources quality assessment includes assessing the percentage of the number of lecturers based on the last education either S1, S2, or S3, the percentage of the number of lecturers based on their position, and the ratio of students to lecturers according to regulations. Institutional Quality Assessment includes accreditation of institutions and study programs, the number of national and international BAN-PT accredited study programs, the number of foreign students and the number of university collaborations. The assessment of Student Activities only includes student performance. Assessment of Research and Community Service Activities includes assessment of research performance, community service performance, and the number of scientific articles (journals and proceedings) published on a local, national, and international scale and those that are not indexed or indexed reputable (Scopus, Thomson Reuters, Copernicus, etc.) per number of lecturers. While the Innovation Quality assessment includes innovation performance (Kemenristekdikti, 2018).

Multi Attribute Utility Theory (MAUT) is one of the decision support system (DSS) methods that can be used to calculate campus rankings. This method uses a scheme in which the final evaluation of an object is defined as a weight summed with a value relevant to its dimensional value (Perdana & Budiman, 2020). Other related studies using DSS method that have been carried out previously to rank universities include the application of the Technique For Others Reference by Similarity to Ideal Solution (TOPSIS) method in ranking

*Putri Maulidina Fadilah

campuses in Medan (Lestari & Mardiana, 2020), the use of the Višekriterijumsko Compromise Rangiranje method (VIKOR) to rank universities in Medan (Perdana & Budiman, 2021), the application of weighting modifications using the Majority Vote method contained in the Bagging method in the Weihgted Product (WP) method (Perdana & Farhana, 2022), and the application of the Majority Vote (MV) method in modifying the weighting of the Multi Attribute Utility Theory (MAUT) method in the same case, campus ranking in Medan (Farhana, Perdana, & Fadilah, 2022). The results of previous studies obtain the optimal results in the highest ranking of several universities in Medan.

However, the determination of the weight of the MAUT method is subjective based on the expertise of the researcher. The objective weighting calculations are needed so that the final results obtained also become more objective. Weight of Evidence (WoE) can be used in feature selection as one method to evaluate the importance of independent variables in predicting target variables. WoE assigns a numerical score to each category of independent variables that describe the strength of its relationship with the target variable (Batar & Watanabe, 2021; Lund, 2021; Parsai & Kumar, 2021). For selecting the most informative and relevant independent variable in predicting the target variable, Information Value (IV) can be used (Borgonovo, Hazen, Jose, & Plischke, 2021; Gupta, Kumar, Kaur, & Tandon, 2022; Verma, 2020; Wang et al., 2020). In this article, the author will examine the determination of the weighting MAUT using WoE and IV methods to rank several universities in Medan. Basically, this research is the next research from the research by one of the authors who explained the problem of campus ranking in Medan.

LITERATURE REVIEW

Multi-Attribute Utility Theory (MAUT) is a method of evaluation scheme, where $v(x)$ is an object and x is defined as a weight added with a value relevant to the dimension value. The dimension value in question is the utility value. The MAUT method is used to convert multiple values into numeric values on a scale of $0 – 1$ with 0 representing the worst choice and 1 representing the best value. The results of the ranking order show the quality of an object. The equation in determining the value of $v(x)$ is as follows (Hadinata, 2018).

$$V(x) = \sum_{i=1}^{n} Wi \cdot Vi(x) \qquad (1)$$

Where, $Vi(x)$ is the evaluation value of an i–th object. Wi is the weight that determines the value of how important the i–th element is to other elements. While n is the number of elements. The total weight is 1 (Hadinata, 2018).

The steps in MAUT method are breaking a decision into different dimensions, determining the relative weight on each dimension, listing all alternatives, calculating the value of the Utility normalization matrix for each alternative according to its attributes.

$$U(x) = \frac{\left(x - X_i\right)}{x_i^+ - x_i^-} \qquad (2)$$

Where $U(x)$ is normalized alternative weight, x_i^- minimum criteria value (the worst weight), x_i^+ is the maximum criteria value (the best weight), and x is the alternative weight. The last step is multiplying the utility with the weight to obtain each alternative values.

In this paper, the weight Wi will be replaced by the weight obtained from the WoE and IV methods. Weight of Evidence (WoE) is used to encode an independent variable into a numerical score that describes the strength of its relationship to target variable. WoE helps to understand the effect of categories in target behavior. WoE can be calculated using the following formula (Batar & Watanabe, 2021; Lund, 2021; Parsai & Kumar, 2021).

$$WoE = \ln\left(\frac{number\ of\ event\ in\ the\ class}{number\ of\ total\ event}\right) - \left(\frac{number\ of\ non\ event\ in\ the\ class}{number\ of\ total\ non\ event}\right) \qquad (3)$$

or,

$$Woe = \ln\left(\frac{\%\ event\ in\ the\ class}{\%\ non\ event\ in\ the\ class}\right) \qquad (4)$$

Information value (IV) is one of the most useful technique to select important variables in a predictive model. It helps to rank variables on the basis of their importance. The IV is calculated using the following formula (Borgonovo et al., 2021; Gupta et al., 2022; Verma, 2020; Wang et al., 2020).
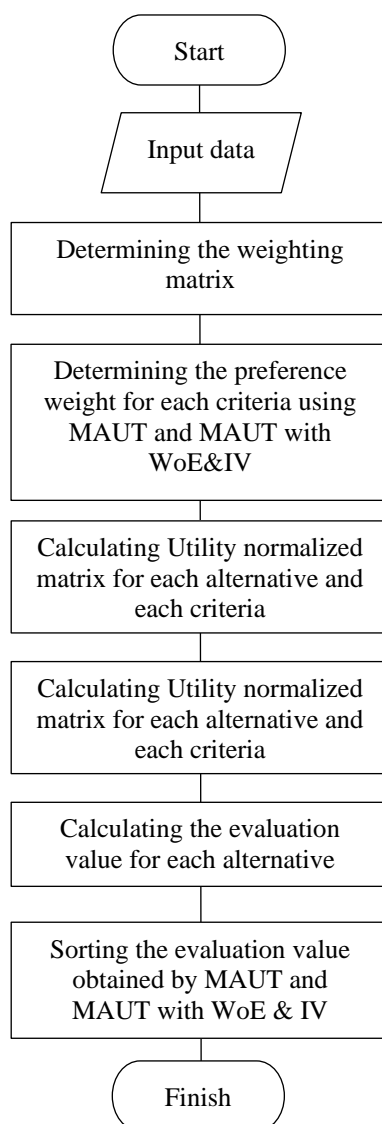
*Putri Maulidina Fadilah

$$IV = WoE \times (\text{ proportion of all events in the class} - \text{proportion of all non events in the class}) \quad (5)$$

By looking at the value of IV, you will get the order of the most important to the least important variables. The use of the IV value in this study is a substitute for the Wj weight in the MAUT formula.

METHOD

The subject in this paper is 10 universities in Medan, North Sumatera, Indonesia (under LLDikti zone 1). The names of the universities are initialized using alphabetical order to avoid conflict and obey the ethics code. Data from the previous research are used in this research in order to compare the MAUT method (Perdana & Budiman, 2020) and MAUT with WoE and IV. The data variables are human resources (HR) data such as the number of lecturers based on their last education (S1, S2, and S3) and the percentage of comparison between the number of lecturers and students, Universities accreditation data, student activity data for each universities, research institute data, community service data, publication data based on sinta classification, and innovation data owned by each university.

```
                    ┌─────────┐
                    │  Start  │
                    └─────────┘
                         │
                   ╱ Input data ╱
                         │
          ┌──────────────────────────────┐
          │  Determining the weighting    │
          │           matrix              │
          └──────────────────────────────┘
                         │
          ┌──────────────────────────────┐
          │  Determining the preference   │
          │  weight for each criteria using│
          │    MAUT and MAUT with         │
          │          WoE&IV               │
          └──────────────────────────────┘
                         │
          ┌──────────────────────────────┐
          │  Calculating Utility normalized│
          │  matrix for each alternative and│
          │         each criteria          │
          └──────────────────────────────┘
                         │
          ┌──────────────────────────────┐
          │  Calculating Utility normalized│
          │  matrix for each alternative and│
          │         each criteria          │
          └──────────────────────────────┘
                         │
          ┌──────────────────────────────┐
          │  Calculating the evaluation    │
          │   value for each alternative   │
          └──────────────────────────────┘
                         │
          ┌──────────────────────────────┐
          │  Sorting the evaluation value  │
          │    obtained by MAUT and        │
          │   MAUT with WoE & IV           │
          └──────────────────────────────┘
                         │
                    ┌─────────┐
                    │ Finish  │
                    └─────────┘
```

RESULT

This study used data from 10 universities in Medan that had been collected in previous studies through observation techniques. The data and criteria used can be seen in the following table.

*Putri Maulidina Fadilah

Table 1
Data

University	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
A	50,8	1	201	32	B	0,074	Madya	Memuaskan	288	0
B	42	18	450	70	A	0,708	Utama	Sangat Bagus	231	0
C	15	8	293	41	B	0	Madya	Memuaskan	110	0
D	17,7	14	166	7	C	0,226	Madya	Memuaskan	346	0
E	33,4	9	148	13	B	0,163	Binaan	Memuaskan	411	0
F	64,6	4	86	14	C	0,034	Madya	Memuaskan	362	0
G	45	10	289	32	B	0,129	Madya	Memuaskan	214	0
H	37,7	46	265	50	B	0	Binaan	Kurang Memuaskan	195	0
I	34	2	223	25	B	0,094	Madya	Memuaskan	242	0
J	1,1	3	57	1	-	0	Binaan	Kurang Memuaskan	1302	0

Notes:
C1 : Percentage the number of lecturers and students
C2 : Lecturer with the latest education S1
C3 : Lecturer with the latest education S2
C4 : Lecturer with the latest education S3
C5 : Campus accreditation
C6 : Student activity
C7 : Research institute (institutional status)
C8 : Community service institution (institutional status)
C9 : Publication (Sinta Score)
C10 : Innovation

From data above, then determine the weights for each criterion, C1 to C10. The weighting that has been done in previous studies is subjective because it is determined by the researcher with consideration of the references. By utilizing the calculation of the WoE and IV methods, a more objective weight is obtained. The results of the calculation of weights for both methods are presented in the following table.

Table 2
Relative Weight by MAUT and MAUT with WoE & IV

Metode	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
MAUT	0,111	0,111	0,111	0,111	0,148	0,037	0,111	0,111	0,111	0,037
MAUT with WoE & IV	0,051	0,051	0,000	0,186	0,097	0,000	0,139	0,139	0,337	0,000

After the relative weight of each criterion are obtained, then list all of alternatives and their weights that have been calculated in previous studies. The result of alternative weight matrix is presented in the following table.

Table 3
Alternative Weight

Alternative	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
A	1	3	4	4	3	2	2	2	4	1
B	1	2	4	4	4	2	3	3	4	1
C	4	2	4	4	3	1	2	2	4	1
D	4	2	4	2	2	2	2	2	3	1
E	2	2	4	3	3	2	1	2	3	1
F	1	2	4	4	2	2	2	2	3	1

*Putri Maulidina Fadilah
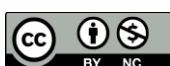
Alternative	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
G	1	2	4	3	3	2	2	2	4	1
H	1	1	3	4	3	1	1	1	4	1
I	2	3	4	4	3	2	2	2	4	1
J	1	2	4	1	1	1	1	1	1	1

Then, calculate the normalized matrix based on the alternative weight matrix contained in Table 3 and the following results are obtained.

Table 4
Normalized Matrix

Alternative	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
A	0	1	1	1	0,67	1	0,5	0,5	1	0
B	0	0,5	1	1	1	1	1	1	1	0
C	1	0,5	1	1	0,67	0	0,5	0,5	1	0
D	1	0,5	1	0,33	0,33	1	0,5	0,5	0,67	0
E	0,33	0,5	1	0,67	0,67	1	0	0,5	0,67	0
F	0	0,5	1	1	0,33	1	0,5	0,5	0,67	0
G	0	0,5	1	0,67	0,67	1	0,5	0,5	1	0
H	0	0	0	1	0,67	0	0	0	1	0
I	0,33	1	1	1	0,67	1	0,5	0,5	1	0
J	0	0,5	1	0	0	0	0	0	0	0

After obtaining the normalized matrix values for each alternative on criterion C1 to C10, the evaluation value of each alternative can then be calculated by multiplying the utility value on the normalized matrix by its criterion weight. The evaluation values of each alternative are presented in the following table.

Table 5
Evaluation Value

Alternative	MAUT	MAUT with WoE & IV
A	0,691358025	0,777541813
B	0,796296296	0,923794719
C	0,709876543	0,802943574
D	0,586419753	0,534414876
E	0,543209877	0,525256324
F	0,549382716	0,607353843
G	0,598765432	0,690268809
H	0,320987654	0,587527994
I	0,728395062	0,79447632
J	0,166666667	0,02540176

Based on the evaluation value in Table 5, although the two methods produce different evaluation values, the highest evaluation value of the two methods is obtained by university B. The MAUT method has an evaluation value of 0.796296296 and the MAUT method with WoE & IV has 0.923794719 for university B. The lowest evaluation value of the two methods was obtained by the same university, university J, with an evaluation value of 0.1666666667 for the MAUT method and 0.02540176 for the MAUT method with WoE & IV. As for the second to ninth order, each method has a different order. For the order of universities from the most popular based on the evaluation value of each method can be seen in the following table.

*Putri Maulidina Fadilah

Table 6
The Rangking of 10 Universities in Medan

Rank	MAUT	MAUT with WoE & IV
1	B	B
2	I	C
3	C	I
4	A	A
5	G	G
6	D	F
7	F	H
8	E	D
9	H	E
10	J	J

DISCUSSIONS

From the result above, the evaluation value by MAUT with WoE & IV for university B is near to 1 than unmodified MAUT. It means that MAUT method with WoE and IV for the weighting, give more optimal performance than MAUT method. There is one thing must be notice, in order to be more sure of optimal results when comparing several methods, we must calculate the error of the methods to see which method is better. So that, the result that we got can be more powerful.

CONCLUSION

Based on the research, MAUT method with the weights using WoE and IV produces a more objective weight so that, the results obtained are also more objective than the ordinary MAUT method. The evaluation value of university B is the highest value for both methods with evaluation value of 0.796296296 using the MAUT method and 0.923794719 using MAUT method with WoE & IV. It means that university B is the most popular university out of ten universities in Medan. Meanwhile, the J university gets the last position with the lowest evaluation scores, 0.1666666667 for the MAUT method and 0.02540176 for the MAUT method with WoE & IV. By knowing that, hopefully J university can improve their quality of the important components of university assessment in order to become a more qualified university not only in Medan but also in Indonesia.

The weighting of the MAUT method with WoE and IV produces a more optimal evaluation value than the subjective weighting of the MAUT method. For the further research, try to use the other methods to determine the weight of MAUT and compare them. It is better if the comparison among the methods is the error of the methods.

REFERENCES

Batar, A. K., & Watanabe, T. (2021). Landslide susceptibility mapping and assessment using geospatial platforms and weights of evidence (WoE) method in the Indian himalayan region: Recent developments, gaps, and future directions. *ISPRS International Journal of Geo-Information*, *10*(3). https://doi.org/10.3390/ijgi10030114

Borgonovo, E., Hazen, G. B., Jose, V. R. R., & Plischke, E. (2021). Probabilistic sensitivity measures as information value. *European Journal of Operational Research*, *289*(2), 595–610. https://doi.org/10.1016/j.ejor.2020.07.010

Farhana, N. A., Perdana, A., & Fadilah, P. M. (2022). PEMERINGKATAN PERGURUAN TINGGI DI KOTA MEDAN MENGGUNAKAN MAJORITY VOTE DALAM PENENTUAN PEMBOBOTAN MULTI-ATTRIBUTE UTILITY THEORY (MAUT). *DSI (Deli Sains Informatika)*, *2*(1), 1–7.

Gupta, V., Kumar, S., Kaur, R., & Tandon, R. S. (2022). Regional-scale landslide susceptibility assessment for the hilly state of Uttarakhand, NW Himalaya, India. *Journal of Earth System Science*, *131*(1). https://doi.org/10.1007/s12040-021-01746-4

Hadinata, N. (2018). Implementasi Metode Multi Attribute Utility Theory (MAUT) Pada Sistem Pendukung Keputusan dalam Menentukan Penerima Kredit. *Jurnal SISFOKOM*, *07*(September), 87–92.

Kemenristekdikti. (2018). Kemenristekdikti Umumkan Peringkat 100 Besar Perguruan Tinggi Indonesia Non Vokasi Tahun 2018. Retrieved February 12, 2019, from Kemenristekdikti website: https://ristekdikti.go.id/kabar/kemenristekdikti -umumkan-peringkat-100-besar-perguruan- tinggi-

*Putri Maulidina Fadilah

indonesia-non vokasi-tahun-2018/

Lestari, Y. D., & Mardiana, M. (2020). Decision Support System For Determining the Best College High Private Using Topsis Method. *SinkrOn : Jurnal Dan Penelitian Teknik Informatika*, *4*(2), 27–33.

Lund, B. (2021). Weight of Evidence, Dummy Variables, and Degrees Of Freedom. *SAS Global Forum 2021*, (L), 1–18.

Parsai, T., & Kumar, A. (2021). Weight-of-evidence process for assessing human health risk of mixture of metal oxide nanoparticles and corresponding ions in aquatic matrices. *Chemosphere*, *263*, 128289. https://doi.org/10.1016/j.chemosphere.2020.128289

Perdana, A., & Budiman, A. (2020). Analysis of Multi-attribute Utility Theory for College Ranking Decision Making. *SinkrOn : Jurnal Dan Penelitian Teknik Informatika*, *4*(2), 19–26.

Perdana, A., & Budiman, A. (2021). College Ranking Analysis Using VIKOR Method. *Journal of Computer Networks, Architecture and High Permformance Computing*, *3*(2), 241–248.

Perdana, A., & Farhana, N. A. (2022). PENERAPAN MAJORITY VOTE METHOD DALAM PENENTUAN PEMBOBOTAN PADA METODE WEIGHTED PRODUCT (WP) PADA PEMERINGKATAN KAMPUS DI KOTA MEDAN. *DSI (Deli Sains Informatika)*, *2*(1).

Verma, P. (2020). Churn prediction for savings bank customers: A machine learning approach. *Journal of Statistics Applications and Probability*, *9*(3), 535–547. https://doi.org/10.18576/JSAP/090310

Wang, W., Lesner, C., Ran, A., Rukonic, M., Xue, J., & Shiu, E. (2020). Using small business banking data for explainable credit risk scoring. *AAAI 2020 - 34th AAAI Conference on Artificial Intelligence*, 13396–13401. https://doi.org/10.1609/aaai.v34i08.7055

*Putri Maulidina Fadilah