Computational model of student competency analysis in fuzzy topsis method

A Nursikuwagus*, L Melian and D Permatasari
Department of Information System, Universitas Komputer Indonesia, Jl. Dipatiukur No. 112-116 Bandung, 40132, Indonesia

*agusnursikuwagus@email.unikom.ac.id

Abstract. In this research is purposed to predict student level competence in sustainability learning through vocational school. Measuring student competence can use a statistical method or intelligent computing process. One method in intelligent computing process is Fuzzy Topsis. An existing problem in this research is every process in evaluation student competence is calculated by mean measure. Challenging process in evaluation competency is use Fuzzy Topsis method. In enrolment Fuzzy Topsis, some variables are needed such as report, psychotest, and teacher recommendation. Result of this research is level of student competency major in vocational schools. Using the data for 270 students, Precision test has calculated 75.60% and Recall test has 96%. This percentage is gained by calculate confusion matrix that differentiate between competence and not competence student level. Impact of this research is competency process can be efficient and effective on processing by Fuzzy Topsis. Other impact is measuring competency can be simple and dynamic than conventional method that used simply mean measure.

1. Introduction
In intelligence process, one method has been applied is Fuzzy Topsis. Fuzzy Topsis is part of soft computing concept [1]. Processing data in intelligence process is making precision and valid to interpret pattern of context. It is hard to solve the problem using hard computing. Since, hard computing is using traversal method and fire match one by one to gain the solution. It is one reason why hard computing is rarely used in many research.

Challenging application in data mining process is measure student competency level at Vocational School. It is activity to separate student who has competency in appropriate skill. Many processes have succeeded to make decisions using Data Mining method [1]. One of challenging process in data mining is mining information by using Fuzzy Topsis. Many papers have been proposed about Fuzzy Topsis. They have been explained about Fuzzy Topsis in several activities. Concept mining in Fuzzy Topsis has proposed by [2-9]. Mahmud Yavuz [10] has explained about applying Fuzzy Topsis in selection equipment problem. In paper [2,8] , many fields have been completed to make solution with Fuzzy Topsis. However, in [2-8] Fuzzy Topsis is just used three to four parameters. The result is not applied in complex parameter yet. Indeed in measuring student competency at vocational school is not solved yet.
Following research that have completed [1-8], this research is directed to make solve a problem about measuring student competence level. Student competency is formulated by Fuzzy Topsis method. Result of this research is provided another measure which purposed measurement of 270 student competency in Fuzzy Topsis [1]. This research has been also provided about Precision and Recall in Fuzzy Topsis.

2. Methodology
2.1. Fuzzy topsis method
Fuzzy Topsis method has proposed by Hwan and Yoon [6] to solve problem in MCDM (Multiple Objective Decision Making) process. Fuzzy Topsis is ideal choice solution. Fuzzy Topsis has had simple formal mathematics concepts, and strategy to accomplish their process. This strategy is directed to step forward as follows normalized matrix decision, weighted normalized matrix, gain positive and negative solution matrix, distance between positive and negative solution, and gain preference value for each alternative. Process in Fuzzy Topsis is needed to calculate performance in rating for each alternative and criteria.

2.1.1. Created Normalized Matrix or Relation Matrix. Normalized Matrix is value that ordered in row and column. Normalized matrix is composed from relation between parameters therein. We have constructed normalize matrix by formula (1) [6].

\[ r_{ij} = \frac{x_{ij}}{\sum_{j=1}^{n} x_{ij}^2} \quad ; \quad i=1,2,...,m \quad \text{and} \quad j=1,2,...,n \]  

(1)

r is item in row i and column j ; x is every value in parameters that composed by rows 1 to i and columns 1 to j [2-8].

2.1.2 Gaining Weighted Normalized Matrix. Weighted normalized matrix is an integer value to give weight for each parameter [6]. This design is tailored by human perception about concern for each parameter. An example, if the research design has five parameters, we can write value for weighted at formula (2) [6].

\[ w_{[ij]} = w_{[ii]} x_{[ii]} r_{[ii]} \quad ; \quad i=1,2,...,m \quad ; \quad j=1,2,...,n \]  

(2)

2.1.3. Gaining Positive and Negative Solution Matrix. Step forward after obtain weighted value is enrolling positive and negative solution matrix. We have created a notation to separate value between positive solution \((A^+)\) and negative solution \((A^-)\). In this step, Fuzzy Topsis is used normalize weighted matrix to make rating \((Y_{ij})\). This formula can be written at formula (3), (4), (5), (6), and (7) [6].

\[ y_{ij} = w_{[ij]} r_{[ij]} \quad ; \quad \text{With} \quad i=1,2,...,m \quad \text{and} \quad j=1,2,...,n \]  

(3)

\[ A^+ = (y^+_1, y^+_2, ..., y^+_n) \]  

(4)

\[ A^- = (y^-_1, y^-_2, ..., y^-_n) \]  

(5)

\[ y^+_i = m \quad \sum_{j=1}^{a} y_{ij} \quad ; \quad i \quad j \quad i. \quad b \]  

(6)

\[ y^-_i = m \quad \sum_{j=1}^{c} y_{ij} \quad ; \quad i \quad i. \quad c \quad a \]  

(7)

2.1.4. Distance between Positive and Negative Solution. Solution for each rating can be remarked between alternative and ideal solution. Formulating the distance can be written at formula (8) and (9) [6].
Distance between $A_i$ and positive ideal solution $D_i^+$:

$$D_i^+ = \left[ \sum_{j=1}^{n} y_i^j - y_i^j \right]^\frac{1}{2}; \quad i = 1, 2, \ldots, m$$  \hspace{1cm} (8)

Distance between $A_i$ and negative ideal solution $D_i^-$:

$$D_i^- = \left[ \sum_{j=1}^{n} y_i^j - y_i^j \right]^\frac{1}{2}; \quad i = 1, 2, \ldots, m$$  \hspace{1cm} (9)

2.1.5. Gaining preference value for each alternative. In this step forward after distance measuring, we are calculated preference value for each ($V_i$). Formulating preference can be written at formula (10) [6].

$$V_i = D_i^- [D_i^- + D_i^+]^{-1}; \quad i = 1, 2, \ldots, m$$  \hspace{1cm} (10)

Rating ($V_i$) is shown as a high competency in major alternative selected ($A_i$) that have a higher preference than the other.

2.2. Precision and recall

Step forward from gaining rating is validation of the formula that is suitable for the case. We have used Precision and Recall method that has significant to problem. Precision and Recall that have used are defined from [7]. We have used formula (11) and (12).

$$R = \frac{T}{T + F}$$  \hspace{1cm} (11)

$$P = \frac{T}{T + F}$$  \hspace{1cm} (12)

Calculating Precision and Recall can be composed by confusion matrix. Confusion matrix content is value from how many result is corrected with real condition. Matrix can be constructed at Table 1.

### Table 1. Composition confusion matrix.

| Condition : A | Not A |
|---------------|-------|
| Test says accepted A | True positive (TP) | False positive (FP) |
| Test says accepted not A | False negative (FN) | True negative (TN) |

3. Results and discussion

3.1. Results

In this research, we have used 270 rows of data [1]. Data have been taken from observed into the field [1]. At vocational school in Indonesia, every school has regulation to accept their student. Prerequisite has to be achieving from every student and accepting students should have to pass examine. There are several parameters that have fulfilled by candidate [10]. These parameters are determined from the school regulation such as national exam, psychology test, interview, grade report, body test, competency test. These parameters that have calculated are executed by Fuzzy Topsis [11].

3.1.1. Dataset.

Following data record can be seen as dataset that have taken from observation. Showing a process, we have shown three dataset as sample. At Table 2 shows an example dataset with parameters therein in Table 2.
Table 2. Example dataset that have gained from survey at Vocational School.

| No | Name           | National Exam | Competency | Grade Report | Body Test | Interview Test | Psychology Test |
|----|----------------|---------------|------------|--------------|-----------|----------------|-----------------|
| 1  | Dadang Wasisto | 30.74         | 83.33      | 79.84        | 80.00     | 85.00          | 50.00           |
| 2  | Tri Andi Kusumah | 28.21       | 73.33      | 79.52        | 80.00     | 85.00          | 80.00           |
| 3  | Ani Suryani    | 31.66         | 83.33      | 79.16        | 70.00     | 65.00          | 80.00           |

3.1.2. Normalized matrix. Step first in Fuzzy Topsis is gained normalize matrix. At Table 3, we are used formula (1) to gain relation matrix or normalized matrix in Table 3.

Table 3. A result from calculating normalized matrix by formula (1) as relation matrix \( r_{ij} \).

| No | Name           | \( x_1 \) | \( x_2 \) | \( x_3 \) | \( x_4 \) | \( x_5 \) | \( x_6 \) |
|----|----------------|-----------|-----------|-----------|-----------|-----------|-----------|
| 1  | Dadang Wasisto | 2.04      | 5.73      | 4.91      | 5.12      | 5.73      | 2.10      |
| 2  | Tri Andi Kusumah | 1.72      | 4.44      | 4.87      | 5.12      | 5.73      | 5.38      |
| 3  | Ani Suryani    | 2.16      | 5.73      | 4.83      | 3.92      | 3.35      | 5.38      |

At Table 3, acquisition of the test score has ran from the use of formula (1). Value in column \( x_1 \), presented by calculating \( x_{11} \) at Table 2 that has divided by the root of the number of values in column \( x_1 \). Result for \( r_{11} \) is 2.04. An example calculation is performed as follows: \( r_{11} = \frac{x_{11}}{\sqrt{\sum_{i=1}^{n} x_{i1}^2}} = \frac{30.74}{\sqrt{30.74^2 + 28.21^2 + \cdots + 33.22^2}} = 2.04 \).

3.1.3. Gaining weighted normalized matrix. Every weight in \([i,j]\) is got from times \( x_{ij} \) and \( r_{ij} \). Following table 4 is an example determines weighted value. At Table 4 is resulted from execution formula (2) in Table 4.

Table 4. An example result for weighted normalized matrix. Y symbol is parameter that has calculated to weight normalized matrix (Matrix \( w \)).

| No | Name           | \( Y_1 \) | \( Y_2 \) | \( Y_3 \) | \( Y_4 \) | \( Y_5 \) | \( Y_6 \) |
|----|----------------|-----------|-----------|-----------|-----------|-----------|-----------|
| 1  | Dadang Wasisto | 62.60     | 477.51    | 392.10    | 409.53    | 487.10    | 105.05    |
| 2  | Tri Andi Kusumah | 48.39     | 325.41    | 387.40    | 409.53    | 487.10    | 430.27    |
| 3  | Ani Suryani    | 68.45     | 477.51    | 382.16    | 274.36    | 217.82    | 430.27    |

At Table 4, the result is gained from calculating weighted with formula (2). As an execution, process is following like: \( w_{11} = x_{11} \times r_{11} = 30.74 \times 2.04 = 62.60 \); for column \( Y_1 \).

We have completed to end of dataset as many as 270 in columns \( Y_1 \). For \( Y_2 \) until \( Y_6 \) is same process like \( Y_1 \).

3.1.4. Gaining Positive and Negative Solution Matrix. Step forward after perform weighted normalize matrix is calculate to solution in positive and negative respectively. We have used formula (3) to collect be \( A^+ \) (alternative positive) and \( A^- \) (alternative negative). \( A^+ \) and \( A^- \), collecting in one matrix, is called \( \begin{bmatrix} y_{ij}^+; y_{ij}^- \end{bmatrix} \). As an execution for \( A^+ \) and \( A^- \), was can be written as follow:

\[
y_1^+ = \max\{\text{column}(Y1)\}; \quad y_2^+ = \max\{\text{column}(Y2)\}; \quad y_3^+ = \max\{\text{column}(Y3)\}.
A^+ = \begin{bmatrix} y_{11}^+; y_{12}^+; \ldots; y_{16}^+ \end{bmatrix} = [79.01; 784.57; 549.01; 583.11; 487.10; 430.27].
\]
$$y^+_i = \min(\text{column}(Y_1)); y^+_j = \min(\text{column}(Y_2)); y^+_k = \min(\text{column}(Y_3))$$
$$A^- = [y^+_{i1}; y^+_{i2}; ...; y^+_{i|\ell|}] = [13.84 ; 22.28 ; 157.91 ; 99.98 ; 217.82 ; 105.05]$$

3.1.5. Distance between Positive and Negative Solution. Distance is counting how far differ between positive and negative solution. Computing distance is used formula (8) and (9). We are proposed D symbol for symbolizing distance, \(D^+\) for solution distance positive, and \(D^-\) for solution distance negative. Running process shows an example execution from formula (8) and (9) in Table 5.

\[
D^+_i = \left[ \sum_{j=1}^{|\ell|} y^+_i - y^-_{ij} \right]^2 ; \quad D^-_i = \left[ \sum_{j=1}^{|\ell|} y^-_{ij} - y^-_{ij} \right]^2
\]

\[
D^+_i = \left( y^+_1 y^+_1 + (y^+_2 - y^-_{12}) + ... + (y^+_6 - y^-_{16}) \right)^2 = (79.01-62.60) + (784.57-477.51) + ... + (430.27-105.05) = 31.29
\]

\[
D^-_i = \left( y^-_{11} y^-_{11} + (y^-_{12} - y^-_{22}) + ... + (y^-_{61} - y^-_{66}) \right)^2 = (62.60-13.84) + (217.82-117.91) + ... + (105.05-105.05) = 8.81
\]

**Table 5.** An Example result positive solution.

| A   | Name             | Y1   | Y2    | Y3   | Y4   | Y5   | Y6   | \(D^+_i\) |
|-----|------------------|------|-------|------|------|------|------|-----------|
| 1   | Dadang Wasisto   | 16.41| 307.06| 156.91| 173.57| 0.00 | 325.22| 31.29     |
| 2   | Tri Andi Kusumah | 30.62| 21.43 | 161.61| 173.57| 0.00 | 19.68 | 8.81      |
| 3   | Ani Suryani      | 10.57| 17.52 | 166.84| 269.28| 8.66 | 27.80 |           |

3.1.6. Gaining preference value for each alternative. Finalization process has determined preference value for each alternative. Preference is gained from formula (10). We have counted every alternative from A1 until A270. An example execution formula (10) can be seen at Table 6.

\[
V_i = D^-_i \left[ D^+_i + D^-_i \right]^{-1} = 8.81 \left[ 31.29 + 8.81 \right]^{-1} = 0.220 ; \quad \text{Preference for alternative A1}
\]

**Table 6.** An Example preference in every alternative A.

| A   | Name             | \(D^+_i\) | \(D^-_i\) | \(V_i\)  |
|-----|------------------|------------|------------|-----------|
| 1   | Dadang Wasisto   | 31.29      | 8.81       | 0.220     |
| 2   | Tri Andi Kusumah | 19.68      | 9.51       | 0.326     |
| 3   | Ani Suryani      | 27.80      | 8.66       | 0.237     |

3.2. Discussion

**Fuzzy Topsis** has a simple execution in process. Every step is just following execution the formula. Subsection before, it has proven that **Fuzzy Topsis** gives simple calculation [2-9]. Values at table have shown strictly computation with numeric manner [11]. At Table 4, information that content at Table 4 is real value. We have taken from survey at Vocational School in Indonesia [1]. We are just to prove steps in **Fuzzy Topsis**. We are thinking is the same in process. 270 dataset or more is the same of process. We are just considering effective algorithm to process big data [8].

Measurement effectively process, we used **Precision** and **Recall**. In formula (11), we have written formulation from [9]. At Table 7 shows a result from inheritance process from calculation every prediction. We have proposed some symbols such as TP (True Positive), FP (False Positive), FN (False Negative), and TN (True Negative). Using confusion matrix is displaying result that classify where the result is same and not as condition. Assisting confusion matrix, it is process for calculating **Precision** and **Recall** like shown at Table 7.
Table 7. Confusion matrix has taken from final result.

|                  | Condition: Accepted | Not Accepted |
|------------------|---------------------|-------------|
| Result accepted  | TP = 192            | FP = 8      |
| Result Not accepted | FN = 62           | TN = 8      |

\[
\text{Precision} = \frac{TP}{TP + FN}^{-1} = 192 \left[\frac{192 + 62}{192 + 8}\right]^{-1} = 0.756 \\
\text{Recall} = \frac{TP}{TP + FP}^{-1} = 192 \left[\frac{192 + 8}{192 + 8}\right]^{-1} = 0.96
\]

In information retrieval, Recall in a process is 0.96. It means that document we are used has relevant 96% in Fuzzy Topsis [7,8]. On the other hand, Precision is 0.756 that is dataset only 75.6% alignments within the context [6-8], that result can be written at Table 8.

Table 8. Result Comparation in Fuzzy Method within Precision and Recall as much as 270 dataset.

| Method          | Precision | Recall |
|-----------------|-----------|--------|
| Fuzzy Mamdani [1] | 75.63%    | 90%    |
| Fuzzy Topsis    | 75.60%    | 96%    |

4. Conclusion
Implementation Fuzzy Topsis in measuring student competency is effective process. Simplicity in execution has made Fuzzy Topsis suitable for dataset that contents numeric values. In parameters therein such as national exam, competency, grade report, body test, interview, and psychology test are made success in predicting student competency. In 270 dataset, we conclude that Fuzzy Topsis is able reach 75.60% in precision. In recall, Fuzzy Topsis reach 96%. Meaning of precision that Fuzzy Topsis can be executed and resulted 75.60% valid data, and meaning of recall that Fuzzy Topsis is able process in document as much as 96% significant document. Comparing with other Fuzzy Method, we have concluded that in Fuzzy term almost the same values in ranging 90% - 96% in Recall. Likewise in precision, the result has the same ranging between 75% - 77%. Measuring competency can be simple and dynamic than conventional method that used only averaging value.

References
[1] Agus N and Agis B 2017 A Mamdani Fuzzy Model to Choose Eligible Student Entry. J.Telkom 15 1 p 365-369
[2] Bazzazi A A, Osanloo M and Soltanmohammadi H 2008 Loading-haulage equipment selection in pen pit mines based on Fuzzy-TOPSIS method J.Min.Res.Man. 24 10 p 87-90
[3] Chen C 2000 Applicability of Fuzzy Topsis Method in Optimal Portfolio Selection and an Application in BIST J.Fuz. Sets.Sys.
[4] Chen C 2000 Extensions of the TOPSIS for Group Decision-Making under Fuzzy Environment J.Fuz. Sets.Sys 1
[5] Haoran S, Kejian L, Haihua P and Yang Y 2016 Fuzzy TOPSIS-Based Supply Chain Optimization of Fresh Agricultural Products J. AMSE 59 1 p 186-190
[6] Hwang C L and Yoon K P 1981 Multiple Attribute Decision Making: Methods and Applications, A State-of-the-Art Survey (Berlin: Springer- Verlang)
[7] Krohling R A and Pacheco A G 2016 A-TOPSIS – An approach Based on TOPSIS for Ranking Evolutionary Algorithms J.Pro.Com.Sci. 308 ScienceDirect
[8] Madi E N, Garibaldi J M and Wagner C 2016 An exploration of issues and limitations in current methods of TOPSIS and fuzzy TOPSIS Int. Conf.on Fuz.Sys. Vancouver, Canada.
[9] Nadaban S, Dzitac S and Dzitac I 2016 Fuzzy TOPSIS: A General View J.Inf.Tech. Quan. Man. 823
[10] Yavuz M 2016 Equipment Selection by using Fuzzy Topsis Method World Mult.Disc.Earth Sci.
[11] Zeng S and Xiao Y 2016 TOPSIS method for intuitionistic fuzzy multiple-criteria decision making and its application to investment selection *J.Kyb.* 45 2 p 282-285