Advances in Social Science, Education and Humanities Research, volume 574
Proceedings of the 6th International Conference on Science, Education and Technology (ISET 2020)

Implementation of Knapsack Problem - Fuzzy Inference System Tsukamoto in the Admission of New Students based on Zone System

Izza Hasanul Muna
Brno University of Technology,
Czech Republic
izza.muna@intermaths.eu

Elok Mutiara Rakhmawati
Universitas Negeri Semarang,
Indonesia
elokmutiara1906@students.unnes.ac.id

***Abstract*---Admission of new students is an annual activity process that exist in educational institutions. In this process, educational institutions try to get the best students among the other candidates according to criteria determined by institutions. In the last 3 years, the Ministry of Education of Indonesia was implementing a new admission system in Elementary, Junior and Senior High School, namely zone system. This paper introduces an approach to optimize admission of new students based on zone system by using Knapsack Problem – Fuzzy Inference System (FIS) Tsukamoto. Specifically, final score was considered as the output for making decision in the admission and 3 variables were used as data input such as: average score of exams, distance, and age. Knapsack problem was used to determine allocation of quota of each region for admission based on average score of exams of all students in a region. Then, data input will be processed by using FIS Tsukamoto to determine the final score of each student. In the last step, rank order list of new students of every region was made. The results show that knapsack problem – FIS Tsukamoto model might be suitable and helpful for proper admission of new students based on zone system.**

Keywords: Knapsack Problem, Fuzzy Inference System, Optimization, Zone System

I. INTRODUCTION

Nowadays, education becomes crucial issue in the world. It is not only a fundamentals human rights that becomes a key to success in the future for everyone, but also it has effect on mind and character of someone. That is why, most of people want to get equality rights in education as the same as the others. This equality is not only limited in learning system, school facilitation or scholarship, but also every aspect in education, including selection system in the admission of new students. Students' admission is defined as annual process compiled systematically to get the best student among the other candidates according to criteria determined by institutions. Overall, in Indonesia, based on Permendikbud No. 44 tahun 2019, there are 4 new student admission pathways, namely zone system, affirmation, the transfer of parents/ guardians, and achievement. The quota allocation for each of pathway is 50%, 15%, 5%, and 30% respectively. It means admission of new students based on zone system is emphasized than the other pathways.

The government made zone system pathway based on the following reasons. First, the government wants to provide a system that reduces segregation in educational institutions that include 3 aspects, socio-economic conditions, learning outcomes and minority group. So, through zone system, equality rights in education can be achieved. Second, the government is not only want to provide a system that provides an opportunity for students accepted at their choices school, but also want to design an effective admission system [1].

After implementing the zone system for 3 years, it is irrefutable that there are some problems need to be solved. Since the zone system prioritizes distance as the main variable for admission of new students, it will affect to an imbalance between the availability of schools and candidate of new students. As the result, several schools will get many students because of its location in dense zone and the other schools will lack of students [1]. Moreover, the zone system with this model will decrease motivation of new students in studying, learning and getting achievement. Previously, most of students study hard and join some additional classes in order to get high score such that favorite school can choose them as new students. Because of zone system, score of exams is not as valuable as before. It can be concluded that admission system of new students based on zone system need to be evaluated.

The study focus on how to optimize admission of new students based on zone system. In this case, knapsack problem and fuzzy inference system (FIS) Tsukamoto are chosen and will be combined together as the method for optimizing the problem. The knapsack problem is one of the most popular combinatorial optimization problems. This problem requires having to fill up a knapsack by selecting from among various possible objects, those which will give the maximum comfort [2]. The knapsack problem can be widely applied in debris collection, resource allocation, job scheduling, capital budgeting, investment decisions, project selection, cargo packing and other fields [3]. In the study, knapsack problem was applied in determining quota allocation for student.

Since the knapsack problem appears in many forms in various integer programming, economics, engineering, and business, this problem have been intensively studied. Vaezi, et.al [4] presented a research about knapsack-based portfolio selection model. Here, expected returns, prices, and budgets are marked with interval values to determine the number of shares of each asset. The result shows that the built model can be used to make decisions in investment. Sampurno, et.al [5] and Al Etawi & Aburomman [6] studied about comparison of dynamic programming algorithm and greedy algorithm on integer knapsack problem. Based on the results, dynamic programming algorithm outperforms greedy algorithm in terms of the optimized solution, while greedy algorithm is better than dynamic programming with respect to runtime and space requirements. In addition, Awasthi & Sharma [7] tried to show convergence on the relative assessment and estimation of dynamic programming, B & B, Greedy and Genetic algorithms. The results of this study are known that Greedy and Genetics algorithm can be used to clear up the 0-1 Knapsack issue inside a sensible time multifaceted nature.

For the second method that will be used in the study is fuzzy inference system (FIS). Here, FIS can be defined as the process of formulating the mapping from a given input to an output using fuzzy logic [8]. Specifically, FIS is a method that interprets the values in the input vector and, based on some sets of rules, assigns values to the output vector [9]. The method generally has four different sections, i.e. a rule base, a fuzzification, an inference engine, and a defuzzification [10]. Mainly, there are three types of FIS: Mamdani, Sugeno, and Tsukamoto [11]. All of them have successfully applied in a wide variety of field, such as automatic control, data classification, decision analysis, expert systems, time series prediction, robotic, and pattern recognition [12]. Therefore, some research related all of FIS method can be found easily.

Alamsyah & Muna [13], Fajrin [14], and Mustika & Sutrisno [15] did research about performance assessment of employee by using FIS Sugeno. The results show that FIS method can be used effectively for making decision regarding performance assessment of employee. Nugraha [16] studied about implementation of FIS Tsukamoto in decision support system of journal acceptance. From the results of this study, it was found that the accuracy of the results of comparison of manual methods, expert decisions, and DSS journal reception using the FIS Tsukamoto was 95% with an error of 5%. This indicates the reception of journals using the Tsukamoto Fuzzy Method is considered accurate and has high precision. In 2012, Alavi [17] did research about FIS Mamdani. The method was applied as a decision making technique to classify the Mozafati dates based on quality. Grading results obtained from the method showed a good general agreement (91%) with the results from the human experts, providing good flexibility in reflecting the expert expectations and grading standards into the results.

In the study, FIS Tsukamoto was chosen for solving the problem because according to [18], the method can define the value that is blurred from the assessment input, build, and apply experiences from experts directly so that they do not go through the training process. Moreover, FIS Tsukamoto can do reasoning on the same principle as human. The reasoning used by FIS Tsukamoto is monotonous reasoning so that its output value can be calculated directly from the membership value associated with its antecedent [12]. Based on the above description, the purpose of this study is to implement knapsack problem - FIS Tsukamoto in the admission of new students based on zone system.

II. METHODS

Quantitative descriptive analysis method is used as the research methodology in the study conducted in several stages described as follows:

ANALYSIS OF NEEDS AND DATA COLLECTION

Before collecting data, determining input and output variables must be done such that data collection process can be done effectively. Overall, there are 4 and 2 variables for input and output respectively. These variables divided into 2 categories according to the process described in Table 1. Data collection process is done by taking the data directly from database of Senior High School (SMAN) 12 Semarang. In the admission of this year, there are 575 applicants who register via zone system and 218 applicants will be selected as candidate of new students of SMAN 12 Semarang.

Table 1. Input and output Variables

Process	Variables	Function
Knapsack Problem	Average score of exams in a region	Input
	Quota allocation	Output
FIS Tsukamoto	Average score of exams	Input
	distance	Input
	age	Input
	Final Score	Output

DATA ANALYSIS TECHNIQUES

In order to apply knapsack problem – FIS Tsukamoto in the study, it is required to have data analysis explained as follows:

DETERMINING QUOTA FOR EACH REGION

Here, it is assumed that to determine quota for each region is similar as how to solve knapsack problem. Thus, the method for solving knapsack problem is performed at this step. Technically, the problem can be formulated as follows:

$$\text{maximize } Z = \frac{\sum_{j=1}^{n} p_j . x_j}{\sum_{j=1}^{n} x_j}$$

subject to:

$$\sum_{j=1}^{n} x_j \leq c,$$

$$x_j > 0, x_j \text{ integer}, j = 1, \ldots n$$

where:
p_j: The score average of applicants in region j.
x_j: Quota for new students in region j.
c: Total capacity for admission of new students.
n: Total region

In order to solve the problem, dynamic programming is chosen since according to [5], dynamic programming algorithm has better performance in terms of the optimal solution than Greedy algorithm. Dynamic Programming is a mathematical technique used to optimize the decision-making process in phases. The optimal decision in all phases is called optimal policy [19]. The pseudo code of the algorithm for the problem is shown on Figure 1.

```
Input:
Array average score for each region (p)
Array 1
The number of region (n)
Total Capacity for admission (c)
for i from 1 to n do
      for j from 0 to c do
            If 1 <= j then
            m[i, j]:= max(m[i-1, j], m[i-1, j] + p[i])
            //student in region n can be included in
            the optimal solution
            //store the optimal solution
            else
             m[i, j]:= m[i-1, j]
            //student in region n cannot be included in
            the optimal solution
            end if
      end for
end for
```

Figure 1. Pseudo code Dynamic Programming Algorithm

FUZZIFICATION

As the first step in FIS method, fuzzification tries to determine the degree of membership in all fuzzy sets of each variable using membership functions. First of all, define the fuzzy inputs and its domain. These inputs are determined by heuristic knowledge and somehow subjective. Table 2 represents fuzzy inputs that affect to membership functions.

Table 2. Fuzzy inputs of each variable

Input Variable	Fuzzy Input sets Category	Domain	Type of Membership
Distance	short	[0,3,6]	Trapezoidal
	average	[3,6,8]	Triangular
	far	[6,8,>8]	Trapezoidal
Average score of exam	Low	[0,60,75]	Trapezoidal
	Mid	[60,75,85]	Triangular
	high	[75,85,100]	Trapezoidal
Age	young	[<16,16,18]	Trapezoidal
	mid	[16,18,20]	Triangular
	old	[18,20,21]	Trapezoidal

Based on Table 2, the membership functions of each input variable (average score of exams, distance and age) can be illustrated in figure (2), (3) and (4) respectively. Moreover, as can be seen in Table 2, the membership functions of the problem divided into 2 piecewise linear functions: triangular (tri) and trapezoidal (trapz) functions. Both of them can be mathematically represented as displayed in equation (1) and (2). For the output variable (final score), there are 2 possibilities predicate: *bad* and *good* with its range domain between 0 and 100. These predicate will be used in the next step.

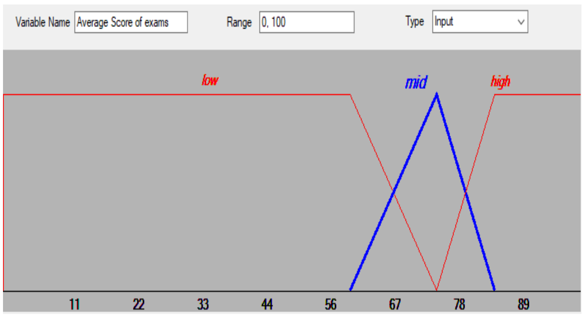

Figure 2. Fuzzy membership function for input average score of exam.

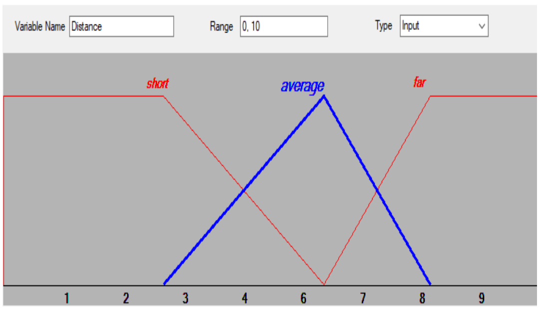

Figure 3. Fuzzy membership function for input **distance**

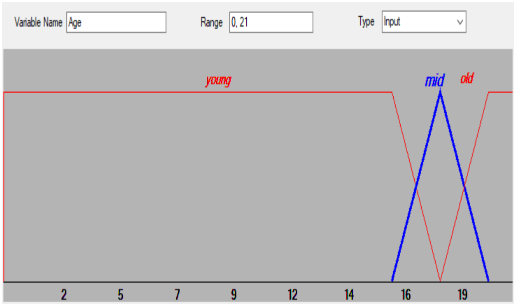

Figure 4. Fuzzy membership function for input age

$$\mu_{tri}(x, a_1, a_2, a_3) = \{0, x \leq a_1 \; or \; x \geq a_2 \; \frac{x-a_1}{a_2-a_1}, a_1 < x \leq a_2 \; \frac{a_2-x}{a_2-a_3}, a_2 < x \leq a_3 \quad (1)$$

$$\mu_{trapz}(x, a_1, a_2, a_3) = \{1, x \leq a_1 \; or \; x \geq a_2 \; \frac{x-a_1}{a_2-a_1}, a_1 < x \leq a_2 \; \frac{a_3-x}{a_3-a_2}, a_2 < x \leq a_3 \; 0, otherwise \quad (2)$$

MAKING INFERENCE RULES

In the study, there will be 27 rules. These rules follow the logic or heuristic knowledge that is derived from the combination of 3 fuzzy sets and 3 input variables. Then method of minimum will be implemented in every rule. An example of rules that has made is given as follows:

[8] If (Distance is short) and (Average score of exams is high) and (Age is young), then (Final Score is Good).

DEFUZZIFICATION

Based on [20], there are 5 methods of defuzzification, height method, centroid method, first (or last) of maxima, mean-max method, and weighted average. Since the research used FIS Tsukamoto, weighted average will be utilized. Equation formula for the method can be represented as follows:

$$z^* = \frac{\sum_{i=1}^{n} \alpha_i . z_i}{\sum_{i=1}^{n} \alpha_i} \quad (3)$$

where:

α_i: The output of i^{th} inference rule

z^*: Final score

Here, z_i represents the output crisp of i^{th} inference rule. In order to compute the value of z_i, consider the domain range of final score. Since the minimum and the maximum value of final score's domain range are 0 and 100 respectively, z_i can be computed by using equation (4) or (5). Now, reconsider the predicate of final score in the previous step. There are 2 predicates: *bad* and *good*. Equation (4) will be used when the i^{th} fuzzy output is *bad* and equation (5) for condition where the i^{th} fuzzy output is *good*.

$$z_i = 100 - (100 \propto_i) \quad (4)$$

$$z_i = (100 \propto_i) \quad (5)$$

CREATING RANK ORDER LIST FOR EACH REGION

This step tries to compile all of output that obtained from knapsack problem and FIS process. The output of knapsack problem is new students' quota for each region, whereas FIS process produces final score as its output. Final score will be utilized for sorting applicants ranking after considering quota of new students in every region. By combining quota allocation and final score of students, top certain number of applicants for each region can be selected as new students.

RESOURCE OPTIMIZER (SOFTWARE SYSTEM DEVELOPMENT)

In order to optimize calculation of the method, a software was developed by using Visual Basic (VB).NET. VB.Net is a Microsoft window programming language that allows the programmer to create and design programs conveniently. In the study, VB.NET was used to develop the software because of its advantages, like supports rich user interfaces (UI) and integrated development environment (IDE), events handling assess to Win 32 API, object-oriented features, error handling, and much more [21]. There are some features in the developed software, such as simulation how the method of knapsack problem – FIS Tsukamoto works, load database and FIS rules, creating rank order list, and export the result of either knapsack problem or FIS Tsukamoto.

III. RESULTS AND DISCUSSION

In this section, detail result and the process of knapsack problem–FIS Tsukamoto method in determining final score is given. First, consider data shown in Table 3. As can be seen in Table 3, the number of total capacity for admission is 218 and there are 4 regions, Ungaran, Gunungpati, Sampangan, and Banyumanik. Each of them has different average score. By using dynamic programming algorithm, it can be easily determined quota allocation of each region. For the result of this process shown in Table 4.

Table 3. The average score of exams for each region

Region	Total Applicants	Average Score

Banyumanik	29	83.932
Gunungpati	364	83.392
Sampangan	77	82.469
Ungaran	105	83.339
Total Capacity	**218**	

Table 4. The result of knapsack problem process

Region	Total Applicants	Quota Allocation
Banyumanik	29	29
Gunungpati	364	189
Sampangan	77	0
Ungaran	105	0

Table 5. Data of Candidate Student

Candidate Name	Distance	Average Score of Exams	Age
Candidate 7	4	82	16

Now, consider data of candidate student in Table 5. This data will be used for illustrating how to process data by using FIS Tsukamoto. First, find the degree of membership of each variable. The result of this stage shown in Table 6. For the next stage is finding α_i (alpha-predicate) and z_i for every inference rule. Here, only 8 rules that affect to the final score since their value of α is not equal to zero. That rules are rule 4, 5, 8, 9, 14, 15, 17 and rule 18. An illustration how to find α and z_i of a rule is given as follows.

[8] If (Distance is short) and (Average score of exams is high) and (Age is young), then (Final Score is Good).

$$\propto_8 = \mu_{distance_{short}} \cap \mu_{score_{high}} \cap \mu_{age_{young}}$$

$$\propto_8 = (0{,}33; 0{,}3; 0{,}33) = 0{,}3$$

$z_8 = 30$

After the involved rule is identified, the last step is calculating the z^* (final score) as the output by using equation (3). As the result, final score of candidate 7 is 39,689. This calculation will be implemented to software such that final score of all of applicants can be found easily.

Table 6. Degree of membership for each variable for candidate 7

Variable	Fuzzy input set category	Degree of membership
Distance	short	0.67
	average	0.33
	far	0
average score of exam	low	0
	mid	0.7
	high	0.3
age	young	0.33
	mid	0.67
	old	0

Figure 5 and 6 show rank order list for region Banyumanik and Gunungpati respectively. These rank order list was made by combining quota allocation and final score of students. While final score utilized for sorting applicants ranking, quota allocation was used to determine top certain number of applicants for every region. Now, consider the result of some candidates shown in Table 7. With pieces such as tables, a bit difficult to determine the final score for each candidate. By using FIS Tsukamoto, the final score can be determined easily. Here, the final score also can represent heuristic knowledge of human that implemented on the system. That is why some results somehow subjective, i.e. candidate 9 and candidate 59. As can be seen in Table 7, the distance and age of candidate 9 less than candidate 59, but the final score of candidate 59 is higher than candidate 9 since the average score of candidate 59 is higher than candidate 9. Based on this comparison, it can be concluded that rule-base in FIS Tsukamoto has important role in determining the final score. Rule – base will determine how good the quality of the output is and how objective the assessment is done.

Table 7. The final score of some candidates of new students

Name	Distance	Score Average	Age	Final Score
candidate 9	1.39	75.58	16	50.4
candidate 59	4.92	89.13	17	45.7
candidate 116	2.73	86.41	15	100
candidate 117	5.73	87.12	15	83.49
candidate 108	0.21	76.46	18	75.03

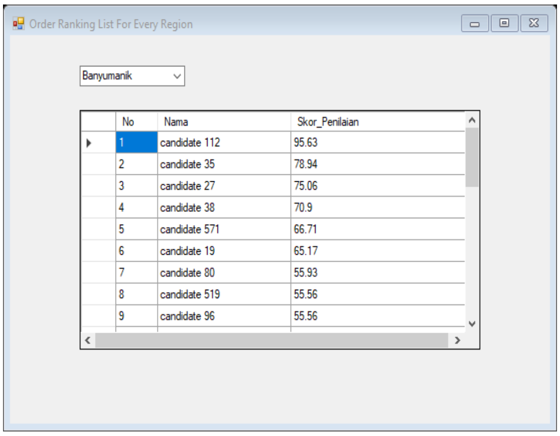

Figure 5. Rank order list of applicants in region Banyumanik

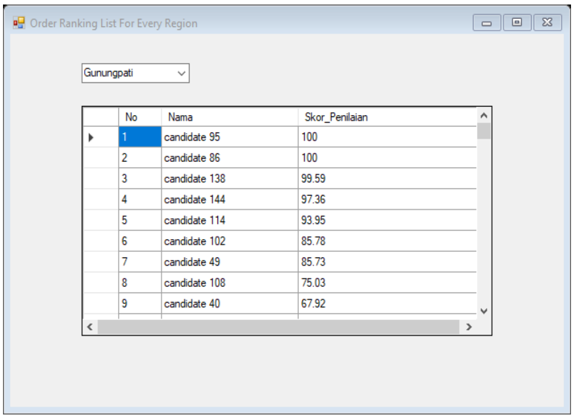

Figure 6. Rank order list of applicants in region Gunungpati

IV. CONCLUSION

In this paper, an approach to optimize admission of new students based on zone system by using Knapsack Problem – Fuzzy Inference System (FIS) Tsukamoto was discussed. The knapsack problem was used to determine quota allocation for each region, whereas FIS Tsukamoto was used to determine the final score. By using FIS Tsukamoto, some results that obtained somehow subjective because of rule-base. Rule – base will determine how good the quality of the output is and how objective the assessment is done. Based on research done by the method, it can be concluded that Knapsack problem – FIS Tsukamoto can be used as an alternative way and might be suitable and helpful for proper admission of new students based on zone system.

REFERENCES

[1] A. Nurlaliliyah, Analisis Kebijakan Sistem Zonasi Terhadap Perilaku Siswa SMP di Yogyakarta, "*Realita: Jurnal Penelitian dan Kebudayaan Islam, 17(1),* pp. 13-21, 2019.

[2] O. K. Peasah, S. K. Amponsah a D. Asamoah, Knapsack problem: A case study of garden city radio (GCR), Kumasi, Ghana, "*African Journal of Mathematics and Computer Science Research, 4(4),* pp. 170-176, 2011.

[3] X. Pan a T. Zhang, Comparison and Analysis of Algorithms for the 0/1 Knapsack, "*IOP Conf. Series: Journal of Physics: Conf. Series, 1069,* pp. 1-7, 2018.

[4] F. Vaezi, S. J. Sadjadi a A. Makui, A portfolio selection model based on the knapsack problem under uncertainty, "*PLOS ONE,* pp. 1-19, 2019.

[5] G. I. Sampurno, E. Sugiharti a A., Comparison of Dynamic Programming Algorithm and Greedy Algorithm on Integer Knapsack Problem in Freight Transportation, "*Scientific Journal of Informatics, 5(1),* pp. 40-50, 2018.

[6] N. A. Al Etawi a T. F. Aburomman, 0/1 Knapsack Problem: Greedy vs Dynamic Programming, "*International Journal of Advanced Engineering and Management Research, 5(2),* pp. 1-10, 2020.

[7] Y. Awasthi a A. Sharma, Contrasting of Various Algorithmic Techniques to Solve Knapsack 0-1 Problem, "*International Journal Journal on Informatics Visualization, 4(1),* pp. 16-21, 2020.

[8] M. Kalinic a M. J. Krisp, Fuzzy inference approach in traffic congestion detection, "*Annals of GIS, 25(4),* pp. 329-336, 2019.

[9] S. Kalogirou, Designing and Modeling Solar Energy Systems, San Diego: Academic Press, 2009, pp. 553-664.

[10] J. Nabati, A. Nezami, E. Neamatollahi a M. Akbari, GIS-based agro-ecological zoning for crop suitability using fuzzy inference system in semi-arid regions, "*Ecological Indicators, 117,* pp. 1-15, 2020.

[11] A. Nazari, F. P. Torgal, A. Cevik a J. G. Sanjayan, Handbook of Alkali-Activated Cements, Mortars and Concretes, Cambridge: Woodhead Publishing, 2015.

[12] M. Sayuti a D. K. Sofyan, Manufacturing optimization using Tsukamoto, "MATEC Web of Conferences 204, pp. 1-9, 2018.

[13] Alamsyah a I. H. Muna, Metode Fuzzy Inference System untuk Penilaian Kinerja, "Scientific Journal of Informatics, 3(1), pp. 88-98, 2016.

[14] A. A. Fajrin, Fuzzy Inference System Sugeno Untuk Evaluasi Kinerja Pelayanan Pegawai Kantor Camat Batam Kota, "Jurnal Positif, 3(2), pp. 83-87, 2017.

[15] F. A. Mustika a Sutrisno, Model Evaluasi Kinerja Karyawan Dengan Metode Fuzzy Sugeno Pada Resto Abtl, "Jurnal String, 1(1), pp. 89-96, 2016.

[16] E. Nugraha, A. P. Wibawa, M. L. Hakim, U. Kholifah, R. H. Dini a M. R. Irwanto,

Implementation of fuzzy tsukamoto method in decision, “Journal of Physics: Conference Series, pp. 1-6, 2018.

[17] N. Alavi, Quality determination of Mozafati dates usingMamdani fuzzy inference system, “Journal of the Saudi Society of Agricultural Sciences, pp. 137-142, 2012.

[18] A. A. Syahidi, F. Biabdillah a F. A. Bachtiar, Perancangan Dan Implementasi Fuzzy Inference System (Fis) Metode Tsukamoto Pada Penentuan Penghuni Asrama, “Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK), 6(1), pp. 55-62, 2019.

[19] D. V. Fawzy, S. Sangadji a S. As'ad, Integer 1/0 Knapsack Problem Dynamic Programming Approach in Building Maintenance Optimization, “International Journal of Science and Applied Science: Conference Series, 2(1), pp. 429-439, 2017.

[20] S. Greenfield a F. Chiclana, Accuracy and complexity evaluation of defuzzification strategies for the discretised interval type-2 fuzzy set, “International Journal of Approximate Reasoning, 54, pp. 1013-1033, 2013.

[21] V. Dwarampudi, S. S. Dhillon, J. Shah, N. S. Sebastian a N. S. Kanigicharla, Comparative study of the Pros and Cons of Programming Languages Java, Scala, C++, Haskell, VB.NET, AspectJ, Perl, Ruby,PHP & Scheme, “Concordia University, Montreal, 2010.