

Performance Evaluation Using the Discrete Choquet Integral: Higher Education Sector

Seher Nur Sülkü¹, Deniz Koçak^{1,*}

¹ Department of Econometrics, Ankara Hacı Bayram Veli University, Ankara, Turkey

ARTICLE HISTORY

Received: 14 November 2018

Accepted: 05 March 2019

KEYWORDS

Performance evaluation,
Fuzzy measure,
Discrete Choquet integral,
k-means

Abstract: Performance evaluation functions as an essential tool for decision makers in the field of measuring and assessing the performance under the multiple evaluation criteria aspect of the systems such as management, economy, and education system. Besides, academic performance evaluation is one of the critical issues in higher institution of learning. Even though the academic evaluation criteria are inherently dependent, most of the traditional evaluation methods take no account of the dependency. Currently, the discrete Choquet integral can be proposed as a useful and effective aggregation operator due to being capable of considering the interactions among the evaluation criteria. In this paper, it is aimed to solve an academic performance evaluation problem of students in a university in Turkey using the discrete Choquet integral with the complexity-based method and the entropy-based method. Moreover, the *k*-means method, which has been widely used for evaluating students' performance over 50 years, is used to compare the effectiveness and the success of two different frameworks based on discrete Choquet integral in the robustness check. Our results indicate that the entropy-based Choquet integral outperforms the complexity-based Choquet and *k*-means method in most of the cases.

1. INTRODUCTION

In recent years, performance evaluation plays an important role due to the lack of operational tools provided objective information in the managerial, educational, and economic areas. Therefore, performance evaluation can be seen as a tool developed for determining whether the wide-ranging set of evaluation criteria is met in the associated areas. Conversely, academic performance evaluation is one of the critical issues in higher institution of learning. Based on this critical issue, many traditional evaluation techniques, which are mainly based on the weighted arithmetic mean, have been widely used, but these techniques only consider situations where all the evaluation criteria are independent. Contrary to the weighted arithmetic mean, the Choquet integral is an appropriate substitute that allows to capture dependency among evaluation criteria (Marichal & Roubens, 2000). The Choquet integral introduced by Choquet

CONTACT: Deniz Koçak ✉ denizkocak36@gmail.com 📍 Department of Econometrics, Ankara Hacı Bayram Veli University, Ankara, Turkey

ISSN-e: 2148-7456 / © IJATE 2019

is an aggregation operator that is extensively employed in quantitative problems such as multi-criteria and multi-objective optimization problems, economics problems, and multi-regression problems, etc. (Choquet 1954; Cui & Li 2008; Angilella et al., 2017). Moreover, the Choquet integral provides an indirect method that reflects the relative importance of evaluation criteria, dependency among them, and their ordered positions in these problems (Angilella et al., 2015; Xu 2010).

Early 2000s, the data mining techniques have been used in the educational area and Educational Data Mining (EDM) has emerged (Baker & Yacef 2009; Peña-Ayala, 2014). In recent years, the tools of the EDM are widely used with educational data (Slater et al., 2017). The new operational tools that serve accountability policies have emerged (Huber & Skedsmo 2016). However, the research in educational data mining have generated the need for rethinking of these new operational tools in handling dependent evaluation criteria. Besides, it is established that more research is needed to specify educational goals for a valid evaluation of students' skills (Herde et al., 2016). In recent years, Shieh, Wu and Liu (2009) proposed discrete Choquet integral with a complexity-based method to evaluate students' performance where the discrete Choquet integral is an adequate aggregation operator which takes the interactions into account. Chang, Liu, Tseng and Chang (2009) found out the poor performance of the traditional regression models in the evaluation of the students' performance when there are interactions among the attributes with using a real data set from a junior high school; and then showed that multiple-mutual information based Choquet integral regression models provide better performance while comparing the joint entropy based and complexity based Choquet integral. In another study, Wang, Nian, Chu and Shi (2012) used the nonlinear multi-regression based on the Choquet integral in order to evaluate the final grade of the students considering previous records such as scores of tests, the average score of quizzes, the number of absent class meeting and the number of incomplete homework as interactive predictive attributes. Branke, Correnre, Greco, Slowinski and Zielniewicz (2016) used Choquet integral as a preference model and suggested an interactive multiobjective evolutionary algorithm.

The discrete Choquet integral has been newly started to be preferred by the researchers due to their success in terms of considering the evaluation criteria dependency. The method is an important kind of non-additive integrals (Wang & Ha 2008), and nowadays its theory is applied by the authors in decision making problems (Grabisch, 1996). Nevertheless, we encountered that there is still a limited number of studies in this context. Only the mentioned studies take the interaction among criteria into account in the literature of academic performance evaluation. Therefore, the purpose of this study is to use various discretization methods and the discrete Choquet integral in order to provide realistic evaluation in educational system. More precisely, the academic performance of students from a university in Turkey are evaluated employing both the entropy-based and the complexity-based discrete Choquet integral and the k -means method. Thereafter, the effectiveness and success of the different discretization techniques are compared, and the model evaluation of these different methods is carried out. The steps of the present analysis are summarized in [Figure 1](#).

In discretization process, a nonoverlapping partition of a continuous domain is obtained. For this aim, first of all continuous attributes are sorted and then the number of intervals are defined. For example, if there will be k intervals then there will be $k-1$ split points. Thus, a researcher actually defines intervals by deciding on the place of split points. Thereafter, all continuous attributes falling into the same interval are automatically mapped to the same categorical value. Hence, the key task is finding meaningful intervals in discretization (Kononenko and Kukar 2007). The equal width interval methods divide the continuous data into the categorical data by using user specified number of intervals. In case of "equal threshold" of the equal width interval methods, if there are $n \times 1$ vectors consisting of three continuous variables, i.e. X , Y and Z , the

data matrix is obtained by assigning the same threshold value to all of them. On the contrary, in case of “not equal threshold”, the data matrix was obtained by assigning a different threshold for X , Y and Z . Then the entropy and complexity based methods are applied to this matrix. The results of these methods are intermingled with the discrete Choquet integral.

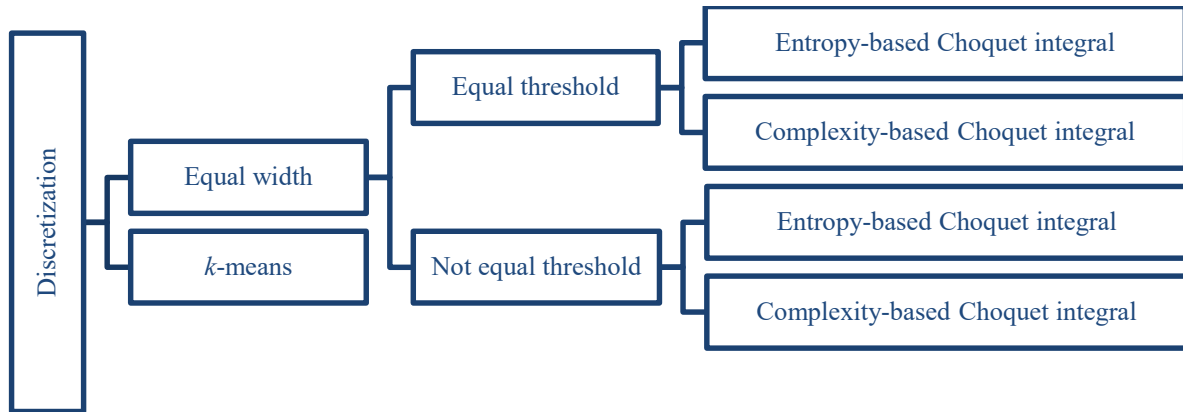


Figure 1. Overview of the discretization methods

Besides, the k -means method is used to compare the effectiveness and the success of two different frameworks based on discrete Choquet integral in the robustness check. Regardless of the fact that the method was presented many years ago, it is one of the most widespread classification algorithms and widely used for evaluating students' performance in educational data mining (Veeramuthu et al., 2014; Jain, 2010). For this reason, the k -means method is not explained technically, but its results in the robustness check is presented.

In this study, the aim is to provide a sufficient and comprehensible background on the discrete Choquet integral method, thus the empirical analysis of the study is exemplified step by step. It is believed that a reader who is even unfamiliar to the Choquet integral methodology can redo the present analyses following the steps which are explained thoroughly in the main text. The rest of this paper is organized as follows. Section 2, a brief introduction of the the discretization techniques, outline of the the fuzzy measure, and the discrete Choquet integrals with entropy-based and complexity-based constructs are presented. The research findings and the robustness check results are presented and discussed in Section 3. Finally, Section 4 concludes the study.

2. METHOD

2.1. Discretization

The evaluation of the academic performance can be considered as a multi-criteria decision making (MCDM) problem. In these problem refers to the evolution of a partition matrix of a data set, and describing the component of a data set from the most preferred alternatives to the least preferred alternatives (Zopounidis & Doumpos, 2002). In many real-life decision making problems that have multi criteria, it is important to preprocess data to effectively apply the algorithms (Kononenko et al., 2007).

Preprocessing the data has a number of steps such as data transformation, cleaning, and data reduction (Pyle, 1999). Currently, discretization is one of the most popular reduction techniques (Garcia et al., 2013). The aim of discretization is to transform continuous attributes which take infinitely many values into categorical attributes and which are significantly reduced subset of discrete values to make the representation of information easier and to learn from the data more accurately and fast (Liu et al., 2002). The discretization methods are summarized in Table 1 (Dougherty et al.,1995).

Detailed review on the discretization methods can be found in Garcia et al., (2013) and Liu et al., (2002). The main separation between discretization methods is whether the class information is employed or not. In the supervised discretization, the class information is considered in the classification but not in unsupervised discretization. Another distinction between discretization methods is global versus local discretization. Global discretization methods use the complete instance space to discretize whereas local discretization methods use only a region of the instance space (Chmielewski & Grzymala-Busse 1996).

The basic unsupervised methods, equal frequency and equal width, do not perform well when there are outliers in the data and when continuous attributes do not follow the uniform distribution (Tan et al., 2005; Catlett, 1991). To deal with these shortcomings, supervised discretization methods have been developed and class information is used to establish the appropriate intervals. There are not as many unsupervised methods as supervised methods, that may be related to the fact that discretization is usually related with the classification task. However, if the class information is not available, only unsupervised methods can be used.

Table 1. Summary of discretization methods

	Global	Local
Supervised	1RD Adaptive Quantizers Chi Merge (Kerber) D-2 (Catlett) Fayyad and Irani / Ting Supervised MCC Predictive Value Max.	Vector Quantization Hierarchical Maximum Entropy Fayyad and Irani C4.5
Unsupervised	Equal width interval Equal frequency interval Unsupervised MCC	<i>k</i> -means clustering

The unsupervised discretization methods can be regarded as sorting problems or separating problems that distinguish the probability occurrences from a mixing of probability laws (Potzelberger & Felsenstein 1993). However, in these methods, the aggregation operators are needed for the fusion of several input values into a single output value (Calvo et al., 2002). In this respect, the discrete Choquet integral is a suitable aggregation operator by taking into the dependency among criteria account (Wen et al., 2016). Besides, the Choquet integral is remarkable in terms of modeling specific interactions of such a broad spectrum of topics including education, health, living conditions (Kasparian & Rolland 2012).

2.2. Fuzzy measure and the discrete Choquet integral

The definitions of fuzzy measures and Choquet integral are as follows (Shieh et al., 2009):

Definition 1. Let N be a finite set of criteria and $P(N)$ be the power set of N . A discrete fuzzy measure (μ) on N is a set function $\mu: 2^N \rightarrow [0,1]$ which satisfies the following axioms. Besides, $\forall S \subseteq N, \mu(S)$ can be explained as the weight of the coalition S .

- (1) $\mu(\emptyset) = 0, \mu(N) = 1$ (boundary condition)
- (2) $A \subseteq B \Rightarrow \mu(A) \leq \mu(B), A, B \in P(N)$ (monotonicity)

Definition 2. Let μ be a fuzzy measure on $N = \{1, 2, \dots, n\}$. The discrete Choquet integral of x in connection with μ is defined as:

$$C_v = \sum_{i=1}^n x_{(i)} [\mu(A_{(i)}) - \mu(A_{(i+1)})], \quad (1)$$

where $(.)$ implies a permutation on N such that $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$. Additionally, $A_{(i)} = \{(i), (i+1), \dots, (n)\}$ and $A_{(n+1)} = \phi$.

There is a need for fuzzy measure to calculate the discrete Choquet integral. In this paper, the complexity based and the entropy based fuzzy measure are qualified to be fuzzy measures. The detailed definition of the measures which needs to gratify the fuzzy measure axioms, is given below:

Definition 3. The complexity C of a discrete random variable N is defined as the function which counts the number of different forms in N . C_1 , is defined as equation (2). $\forall S \subseteq N$, to calculate the complexity of the subsets of criteria of N . Clearly, $C_1(\phi) = 0$ and if $A \subseteq B \Rightarrow C_1(A) \leq C_1(B)$, $A, B \in N$. That is C_1 , is a fuzzy measure.

$$C_1(S) = \frac{C(S)}{C(N)}, \quad (2)$$

Definition 4. Let A be a discrete random variable and p^A be the probability of A , then the entropy of A is defined as:

$$h(A) = - \sum p^A \log_2 p^A, p^A > 0. \quad (3)$$

Let B be a discrete random vector which contains at least two discrete random variables, p^B be the joint probability and $h(B)$ the joint entropy. By using the idea of the joint entropy to calculate the entropy of the subsets of criteria of N , the fuzzy measure (μ_1) is defined as:

$$\mu_1(S) = \frac{h(S)}{h(N)}, \forall S \subseteq N. \quad (4)$$

2.3. Evaluation the performance of the models

Usually practical applications that used the entropy-based and the complexity-based discrete Choquet integral evaluate the performance of the models with a metric called as ‘‘accuracy’’. Furthermore, in the applications of k -means method, the cluster evaluations can be done with the measures of cluster cohesion and cluster separation (Tan et al., 2005). However, when different discretization techniques and their different model evaluation methods are compared, the mean square error (MSE) criteria would be more suitable to choose the best performing one among them (Greene, 2016). In this study, MSE was employed to evaluate alternative models performances. While comparing the models, as MSE gets smaller, the model does better performance. Thus, the model with the smallest MSE value is preferred. Let θ be a parameter and $\hat{\theta}$ an estimator of this parameter, the mean square error of an estimator is defined as below:

$$MSE [\hat{\theta}|\theta] = E [(\theta - \hat{\theta})^2]. \quad (5)$$

3. EMPIRICAL STUDY and RESULTS

The raw data set shown in Table 2 is composed of 33 students' course scores from Econometrics Department at Gazi University. The courses are chosen as follows: Introduction to Statistics and Probability-II (D_1), Microeconomics (D_2), Macroeconomics (D_3), Mathematics-II (D_4), and Econometrics-I ($EKON$).

The $EKON$ scores of the students are set as control group in the analysis because Econometrics-I is a discipline that requires comprehensive knowledge of the other four courses. Besides, the minimum and maximum score for each course are 1 and 100, respectively.

In the empirical study of this paper, it is aimed to estimate the Econometrics scores of the students with using the students' scores of Introduction to Statistics and Probability-II, Microeconomics, Macroeconomics and Mathematics-II courses. For this aim, the discrete Choquet integral was used as an aggregation and estimation operator because of the fact that there are interactions among these four courses. Thereafter, to measure the success of the estimation based on the Choquet integral, the mean square error was computed by using the students' raw scores of Econometrics-I (see Table 2), and the estimation scores (see Table 7).

Table 2. Raw data scores of the students

Student	D_1	D_2	D_3	D_4	$EKON$	Student	D_1	D_2	D_3	D_4	$EKON$
1	55.8	42	52	76	30	18	62.6	66.2	66	90	65.2
2	42	63.8	49	94	38	19	66	28.8	45	78	49.4
3	39.6	45	52	50.2	68	20	68	45	73	100	38.8
4	40.4	42	51	94	61.4	21	60	47.2	51	100	48
5	61.6	54.6	56	86	77.8	22	66	41.8	47	92	44
6	67.8	45	77	100	57	23	79.2	58.2	71	100	66.2
7	36.8	47.2	46	90	36.8	24	29.4	41.8	61	73	37.6
8	52	57	54	87	37.2	25	68.4	50.4	79	92	51.6
9	44.8	45	59	100	59.8	26	53.8	34.8	59	93	51.8
10	32.6	44.8	45	86	32.6	27	74	47.8	74	100	73.2
11	62.2	48	50.4	94	43.2	28	46.2	49.8	47	74	19.8
12	67.4	57.8	55	100	39.2	29	68.2	47.8	59	100	60.6
13	64	52.2	45	49	67.8	30	76.8	72	94	100	83.6
14	54	13.2	62	78	2.8	31	56	31.6	47	90	33.6
15	50.4	22.4	47	74	41.8	32	76.8	55.6	78	96	72
16	67.6	53.6	57	94	50	33	72.8	20.2	53	56.6	75.2
17	63.4	42	45	97	51						

First of all, the descriptive statistics and the normality of the data were checked out. As presented in the Table 3, the average of $EKON$ is 50.46 while the averages of D_1 and D_3 are around 60, the average of D_2 is almost 46. The mathematics course has the highest average, almost 88. Since $n = 33$, Kolmogorov-Smirnov test and Jarque Bera test are appropriate for testing normality. With respect to the Jarque Bera test, the null hypothesis of normality for the distribution of returns is rejected at the significance level of 5% and all variables are not normally distributed. Furthermore, according to Kolmogorov-Smirnov test, D_2 , D_3 and D_4 variables are not normally distributed; D_1 and $EKON$ variables are normally distributed (Asymptotic Significance > 0.05).

Table 3. Results of one-sample Kolmogorov-Smirnov test and Jarque Bera

		D_1	D_2	D_3	D_4	$EKON$
Normal Parameters	Mean	58.38	45.90	57.77	87.39	50.46
	Std. Deviation	13.38	12.67	12.38	14.30	17.87
Most Extreme Differences	Positive	0.08	0.08	0.16	0.19	0.08
	Negative	-0.14	-0.191	-0.15	-0.21	-0.07
Test Statistic		0.14	0.191	0.16	0.21	0.08
Asymptotic Significance (2-tailed)		0.09	0.00	0.04	0.00	0.20
Skewness		-0.54	-0.58	1.16	-1.44	-0.32
Kurtosis		-0.59	0.86	0.84	1.536	0.21
Jarque Bera		19.32	8.13	13.77	14.40	11.26

Before applying the complexity-based and entropy-based methods, the number of the level of score (m) which transforms the continuous raw data into the categorical level of the score should be decided on. This level of score can be defined by the users or can be stated in terms of the interval width for the equal width interval method.

Table 4. Categorical data scores of the students ($m = 3$)

Student	D_1	D_2	D_3	D_4	$EKON$	Student	D_1	D_2	D_3	D_4	$EKON$
1	2	2	1	2	2	18	3	3	2	3	3
2	1	3	1	3	2	19	3	1	1	2	2
3	1	2	1	1	3	20	3	2	2	3	2
4	1	2	1	3	3	21	2	2	1	3	2
5	2	3	1	3	3	22	3	2	1	3	2
6	3	2	2	3	3	23	3	3	2	3	3
7	1	2	1	3	2	24	1	2	1	2	2
8	2	3	1	3	2	25	3	2	3	3	2
9	1	2	1	3	3	26	2	2	1	3	2
10	1	2	1	3	2	27	3	2	2	3	3
11	2	2	1	3	2	28	2	2	1	2	1
12	3	3	1	3	2	29	3	2	1	3	3
13	3	2	1	1	3	30	3	3	3	3	3
14	2	1	2	2	1	31	2	1	1	3	2
15	2	1	1	2	2	32	3	3	3	3	3
16	3	3	1	3	2	33	3	1	1	1	3
17	3	2	1	3	2						

First of all, equal thresholds approach of the equal width interval method was used. The equal width interval method converts the continuous data into the categorical data by employing user specified number of intervals. Here the number of intervals as $m = 2, 3, 4, 5, 6, 7, 8,$ and 9 were specified. Thereafter, the raw data in Table 2 was transformed by using “hist.m” program of Matlab for D_1, D_2, D_3, D_4 and $EKON$ variables when $m = 2, 3, 4, 5, 6, 7, 8,$ and 9 . Later, the complexity-based and entropy-based fuzzy measure were computed at each level of score ($m = 2, 3, 4, 5, 6, 7, 8,$ and 9) with applying the equations (1), (2), and (3) to determine the dependency of the evaluation criteria. Final identified fuzzy measures for each subset were computed by Matlab and showed in Table 5. Before presenting the Table 5, in order to make clear that how the final values are obtained $m=3$ case was provided as an example. Here, how each of the steps was followed when $m=3$ was employed is summarized in the preceding

paragraph. Firstly, continuous raw data scores (in Table 2) were converted into categorical data. When $m=3$ is employed, the categorical data score for each course for each student can be 1, 2 or 3. Table 4 shows the categorical data scores for each criterion transformed from the raw data scores by using “hist.m” program of Matlab.

Furthermore, the histograms of the D_1, D_2, D_3, D_4 and $EKON$ courses when the number of the level score is equal to three, $m=3$, can be seen in Figure 2. For example, for Microeconomics (D_2) course, students with grade in the interval of $[0, 32.8)$ constitute the first category and each observation in this group takes categorical value “1”, students with grade in the interval of $[32.8, 52.4)$ constitute the second category, and each observation in this group takes categorical value “2” and students with grade in the interval of $[52.4, 72)$ constitute the third category and each observation in this group takes categorical value “3”.

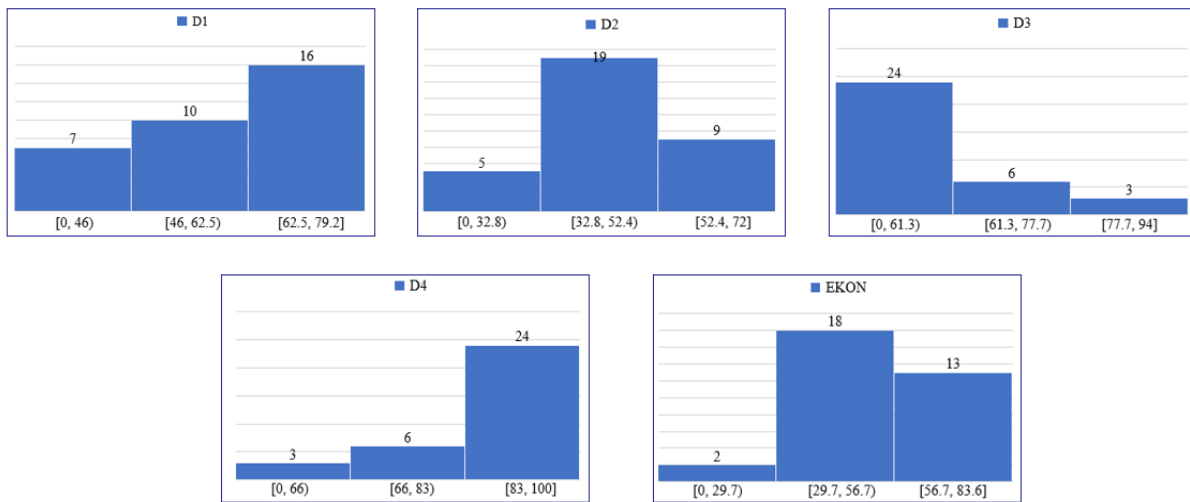


Figure 2. Histograms of the courses for $m = 3$

For instance, in Table 2, the first student’s grade for D_2 is 42, so this student belongs to second category and in Table 4 in the column of D_2 this observation takes value “2”. For each course raw data of grades are converted into categorical data in the same manner. For each histogram of the courses, the first column shows how many times “1” value is repeated, second column shows how many times “2” value is repeated, and the third column shows how many times “3” value is repeated. Besides, the numbers at which intervals correspond to these values are shown below the columns. Now, to obtain entropy based fuzzy measure, $h(N)$ was computed. When the transformed data scores of the students are considered, there are 19 different joint pattern in Table 4 these are: $(2,2,1,2), (1,3,1,3), (1,2,1,1), (1,2,1,3), (2,3,1,3), (3,2,2,3), (2,2,1,3), (3,3,1,3), (3,2,1,1), (2,1,2,2), (2,1,1,2), (3,2,1,3), (3,3,2,3), (3,1,1,2), (1,2,1,2), (3,2,3,3), (3,3,3,3), (2,1,1,3),$ and $(3,1,1,1)$. Besides, how many times the patterns are repeated are given respectively 2, 1, 1, 4, 2, 3, 2, 2, 1, 1, 1, 3, 3, 1, 1, 1, 2, 1, and 1. It means that in Table 4 $(2,2,1,2)$ is repeated twice, $(1,3,1,3)$ is repeated once, and so on. Thus, the joint probabilities are defined and then the entropy of the finite set of criteria (N) employing equation 3 could be calculated as:

$$\begin{aligned}
 h(N) &= - \sum p \log_2 p \\
 &= -0.06 * \log_2(0.06) - 0.03 * \log_2(0.03) - 0.03 * \log_2(0.03) - 0.12 * \log_2(0.12) - 0.06 * \log_2(0.06) - 0.09 * \\
 &\quad \log_2(0.09) - 0.06 * \log_2(0.06) - 0.06 * \log_2(0.06) - 0.03 * \log_2(0.03) - 0.03 * \log_2(0.03) - 0.03 * \\
 &\quad \log_2(0.03) - 0.03 * \log_2(0.03) - 0.09 * \log_2(0.09) - 0.09 * \log_2(0.09) - 0.03 * \log_2(0.03) - 0.03 * \\
 &\quad \log_2(0.03) - 0.03 * \log_2(0.03) - 0.06 * \log_2(0.06) - 0.03 * \log_2(0.03) - 0.03 * \log_2(0.03) \\
 &= 4.07
 \end{aligned}$$

Now the subsets of criteria of N which are $\forall S \subseteq N$ was introduced: empty set, $\{D_1\}$, $\{D_2\}$, $\{D_3\}$, $\{D_4\}$, $\{D_1, D_2\}$, $\{D_1, D_3\}$, $\{D_1, D_4\}$, $\{D_2, D_3\}$, $\{D_2, D_4\}$, $\{D_3, D_4\}$, $\{D_1, D_2, D_3\}$, $\{D_1, D_2, D_4\}$, $\{D_1, D_3, D_4\}$, $\{D_2, D_3, D_4\}$, and $\{D_1, D_2, D_3, D_4\}$. These subsets were symbolized as respectively: (0,0,0,0), (1,0,0,0), (0,1,0,0), (0,0,1,0), (0,0,0,1), (1,1,0,0), (1,0,1,0), (1,0,0,1), (0,1,1,0), (0,1,0,1), (0,0,1,1), (1,1,1,0), (1,1,0,1), (1,0,1,1), (0,1,1,1), and (1,1,1,1) as shown in Table 5. For example, the effect of the only $\{D_1\}$ course is known, that situation is symbolized as (1,0,0,0); when the effect of the $\{D_1, D_2\}$ courses is known, that situation is symbolized as (1,1,0,0). Then the entropy of the subsets of criteria of N , i.e. $h(S)$ is calculated using equation 3. For example in order to calculate $h(D_1)$ Table 4 is considered and the column of D_1 is observed to see how many times “1”, “2” and “3” categories are repeated; “1” is repeated 7 times, “2” is repeated 10 times, and “3” is repeated 16 times. $h(D_1)$ is calculated as follow:

$$h(D_1) = -\frac{7}{33} * \log_2 \left(\frac{7}{33}\right) - \frac{10}{33} * \log_2 \left(\frac{10}{33}\right) - \frac{16}{33} * \log_2 \left(\frac{16}{33}\right) = 1.50$$

For instance, if $h(D_1, D_2)$ is considered, D_1 and D_2 columns are simultaneously examined and it is seen that “2, 2” case appears five times, “1, 3” once, “1, 2” six times and so on, thus:

$$\begin{aligned} h(D_1, D_2) &= -\frac{5}{33} * \log_2 \left(\frac{5}{33}\right) - \frac{1}{33} * \log_2 \left(\frac{1}{33}\right) - \frac{6}{33} * \log_2 \left(\frac{6}{33}\right) - \frac{2}{33} * \log_2 \left(\frac{2}{33}\right) - \frac{8}{33} \\ &\quad * \log_2 \left(\frac{8}{33}\right) - \frac{7}{33} * \log_2 \left(\frac{7}{33}\right) - \frac{3}{33} * \log_2 \left(\frac{3}{33}\right) - \frac{2}{33} * \log_2 \left(\frac{2}{33}\right) \\ &= 2.76 \end{aligned}$$

Thus, the entropies of the selected subsets as an example are calculated as follow:

$$\begin{aligned} h(D_1) &= 1.50 \\ h(D_1, D_2) &= 2.76 \\ h(D_1, D_2, D_3) &= 3.50 \\ h(D_1, D_2, D_3, D_4) &= 4.07 \end{aligned}$$

Now, the fuzzy measures can be obtained by employing equation 4 as $\mu_1(S) = \frac{h(S)}{h(N)}$, ($\forall S \subseteq N$). As shown in Table 5 for $m=3$, the entropy based fuzzy measures for the selected subsets as an example are defined as follows. Besides, the entropy based fuzzy measure of the empty set is always equal to 0.

$$\begin{aligned} \mu_1(D_1) &= \frac{h(D_1)}{h(N)} = \frac{1.50}{4.07} = 0.37 \\ \mu_1(D_1, D_2) &= \frac{h(D_1, D_2)}{h(N)} = \frac{2.76}{4.07} = 0.68 \\ \mu_1(D_1, D_2, D_3) &= \frac{h(D_1, D_2, D_3)}{h(N)} = \frac{3.50}{4.07} = 0.86 \\ \mu_1(D_1, D_2, D_3, D_4) &= \frac{h(D_1, D_2, D_3, D_4)}{h(N)} = \frac{4.07}{4.07} = 1 \end{aligned}$$

The entropy based fuzzy measures for $m=3$ is obtained, and then the complexity based fuzz measures is obtained. Firstly, the complexity of the discrete random variable, i.e. $C(N)$ is needed to be computed in equation 2. When the transformed data scores of the students were considered, there was 19 different joint pattern i.e., (2,2,1,2), (1,3,1,3), (1,2,1,1), (1,2,1,3), (2,3,1,3), (3,2,2,3), (2,2,1,3), (3,3,1,3), (3,2,1,1), (2,1,2,2), (2,1,1,2), (3,2,1,3), (3,3,2,3), (3,1,1,2), (1,2,1,2), (3,2,3,3), (3,3,3,3), (2113), and (3,1,1,1) (see Table 4). Thus, through the complexity counts the number of different pattern is $C(N) = 19$.

Thereafter, the complexity of the subsets of criteria of N , i.e. $C(S)$ is calculated. For instance, there are three features in D_1 : 1,2,3; there are three features in D_2 : 1,2,3; there are three features in D_3 : 1,2,3; thus, the complexities of the selected subsets as an example are calculated as follow:

$$\begin{aligned} C(D_1) &= 3 \\ C(D_1, D_2) &= 8 \\ C(D_1, D_2, D_3) &= 13 \\ C(D_1, D_2, D_3, D_4) &= 19 \end{aligned}$$

Similarly, after the complexity for each subset of N is calculated, the complexity based fuzzy measures can be obtained by employing equation 2 as $C_1(S) = \frac{C(S)}{C(N)}$, ($\forall S \subseteq N$). The complexity based fuzzy measures for the selected subsets as an example are computed for $m=3$ as follows and the results are given in Table 5. Besides, the complexity based fuzzy measure of the empty set is always equal to 0.

$$\begin{aligned} C_1(D_1) &= \frac{C(D_1)}{C(N)} = \frac{3}{19} = 0.16 \\ C_1(D_1, D_2) &= \frac{C(D_1, D_2)}{C(N)} = \frac{8}{19} = 0.42 \\ C_1(D_1, D_2, D_3) &= \frac{C(D_1, D_2, D_3)}{C(N)} = \frac{13}{19} = 0.68 \\ C_1(D_1, D_2, D_3, D_4) &= \frac{C(D_1, D_2, D_3, D_4)}{C(N)} = \frac{19}{19} = 1 \end{aligned}$$

Up to now, how the entropy and complexity based fuzzy measures are achieved for $m=3$ have been explained. These values are computed for each level of score ($m = 2, 3, 4, 5, 6, 7, 8$, and 9) in the same manner. Finally, the identified fuzzy measures for each subset are obtained. For $m=3$, the fuzzy measures are summarized in Table 5.

After all fuzzy measures are identified, and it can be said that the entropy based fuzzy measures are relatively larger than the complexity based fuzzy measures. Furthermore, “not equal thresholds approach” in which the variables can have different thresholds is used. As explained in the methodology section, “histogram function”[†] in Matlab is used as bin width optimization method. When “histogram function” is employed, the threshold numbers of D_1, D_2, D_3 , and D_4 courses were found as 6, 7, 6 and 3, respectively. (For *EKON* course, the threshold number was equal to 9). It is observed that the entropy based fuzzy measures are relatively larger than the complexity based fuzzy measures as seen in Table 6.

[†] The function selects the optimal bin size of a histograms by using automatic binning algorithm such as auto, scott, freedman-diaconis, sturges. These algorithms return bins with a uniform width by showing the underlying shape of the distribution.

Table 5. Identified fuzzy measure for $m = 3$ (Equal thresholds)

D_1	D_2	D_3	D_4	Entropy based fuzzy measure	Complexity based fuzzy measure
0	0	0	0	0	0
1	0	0	0	0.37	0.16
0	1	0	0	0.34	0.16
0	0	1	0	0.27	0.16
0	0	0	1	0.27	0.16
1	1	0	0	0.68	0.42
1	0	1	0	0.58	0.32
1	0	0	1	0.60	0.42
0	1	1	0	0.61	0.42
0	1	0	1	0.55	0.37
0	0	1	1	0.53	0.32
1	1	1	0	0.86	0.68
1	1	0	1	0.83	0.74
1	0	1	1	0.73	0.58
0	1	1	1	0.79	0.63
1	1	1	1	1	1

Table 6. Identified fuzzy measure (Not equal thresholds)

D_1	D_2	D_3	D_4	Entropy based fuzzy measure	Complexity based fuzzy measure
0	0	0	0	0.00	0.00
1	0	0	0	0.51	0.21
0	1	0	0	0.51	0.24
0	0	1	0	0.42	0.21
0	0	0	1	0.23	0.10
1	1	0	0	0.86	0.72
1	0	1	0	0.83	0.62
1	0	0	1	0.68	0.45
0	1	1	0	0.80	0.62
0	1	0	1	0.66	0.41
0	0	1	1	0.60	0.34
1	1	1	0	0.99	0.97
1	1	0	1	0.91	0.83
1	0	1	1	0.90	0.76
0	1	1	1	0.90	0.79
1	1	1	1	1	1

After the fuzzy measures are identified, the results are intermingled with the discrete Choquet integral through equation (1). By this way the scores of students' academic performances for both the entropy based Choquet integral method and the complexity based Choquet integral method was obtained. When equal thresholds are used, these obtained scores are transformed according to m level ($m = 2, 3, 4, 5, 6, 7, 8,$ and 9) for each entropy based Choquet integral method and complexity based Choquet integral method.

Now, let's consider equal threshold approach. For example, the number of the level of score is equal to 3 (i.e. $m=3$), and the fuzzy measure is entropy based fuzzy measure. The raw scores of the first student are 55.8, 42, 52, 76 (see Table 2). First of all, the scores should be ranked from the smallest to the largest, i.e., 42, 52, 55.8, 76. Then, the estimation score is computed by the discrete Choquet integral as follow:

Estimation score

$$\begin{aligned}
 &= 42 * \mu_1(D_2, D_3, D_1, D_4) + (52 - 42) * \mu_1(D_3, D_1, D_4) + (55.8 - 52) * \mu_1(D_1, D_4) \\
 &\quad + (76 - 55.8) * \mu_1(D_4) \\
 &= 42 * 1.00 + (52 - 42) * 0.73 + (55.8 - 52) * 0.60 + (76 - 55.8) * 0.27 \\
 &= 57.03
 \end{aligned}$$

After all estimation scores of the students' academic performances is computed, the estimation scores are transformed to the categorical data by using "hist.m" program of Matlab. Finally, both the estimation scores and the transformed scores are showed in Table 7 for each students.

Table 7. Estimation score and the transformed scores of the students for $m = 3$

Student	Estimation score	Transformed score	Student	Estimation score	Transformed score
1	57.03	1	18	71.81	2
2	63.78	2	19	57.33	1
3	47.11	1	20	71.81	2
4	58.02	1	21	66.12	2
5	66.37	2	22	63.93	2
6	72.81	2	23	78.08	3
7	56.26	1	24	52.61	1
8	63.31	2	25	72.73	2
9	63.43	2	26	60.66	2
10	52.71	1	27	74.03	3
11	65.35	2	28	54.89	1
12	71.82	2	29	70.06	2
13	54.83	1	30	86.26	3
14	51.70	1	31	57.42	1
15	50.21	1	32	76.64	3
16	69.71	2	33	52.40	1
17	69.17	2			

When the complexity based fuzzy measure is used, final transformed scores can be obtained similarly with using the discrete Choquet integral. Now, the evaluation of the performances of each models is required. As explained in section 2.4, mean square errors are used to compare the alternative models performances. In the present study, the *EKON* scores of the students are used as control group, actually these scores are the parameters (θ values) and the obtained results by using the alternative methods are the estimators ($\hat{\theta}$ values). The mean of the squared difference between the parameter and the estimator gives the mean squared error value. The mean square errors are calculated for each method for $\forall m = 2, 3, 4, 5, 6, 7, 8, 9$, and the results are shown in Table 8.

Table 8. MSE results (Equal threshold)

m	Complexity based Choquet	Entropy based Choquet
$m=2$	0.36	0.39
$m=3$	0.91	0.85
$m=4$	1.64	1.06
$m=5$	2.33	2.52
$m=6$	3.33	2.03
$m=7$	5.06	3.15
$m=8$	6.03	4.27
$m=9$	8.00	5.48

Finally, the *MSE* results of the methods are summarized in Table 8. Obviously as shown in table, the complexity-based and the entropy-based Choquet integral have the minimum *MSE* results while the number of the level of score (m) is two. However, a binary transformation is not generally preferred in the higher institution of learning. By using the idea of this, it can be seen that the complexity-based Choquet integral while $m = 3, 4, 5$, and the entropy-based Choquet integral while $m = 3, 4, 5, 6$ have relatively small *MSE*. Thus, $m = 3, 4, 5$ can be regarded as possible candidates that should be used in this part of the study. Namely, it can be said that the obtained *MSE* results by using both entropy and complexity based methods are closer to the scores of control group when the number of the level of score is equal to 3, 4 or 5. It is seen that using “equal threshold” Choquet integral both entropy and complexity based provide better results than “not equal threshold” cases in most of the times. The “not equal thresholds” *MSE* results for the entropy and the complexity based Choquet integral are respectively 2.94 and 1.91.

Robustness Check. The k -means is one of the most well-known statistical methods for determining new structure when investigating data sets (Flynt and Dean 2016). The method is widely used for evaluating students’ performances (Veeramuthu et al. 2014). Now, robustness check was provided by comparing k -means performance with Choquet integral applications. Here the intermediate steps of k -means algorithm were not provided. (However, if requested, corresponding author can provide the all steps of robustness check using k -means method).

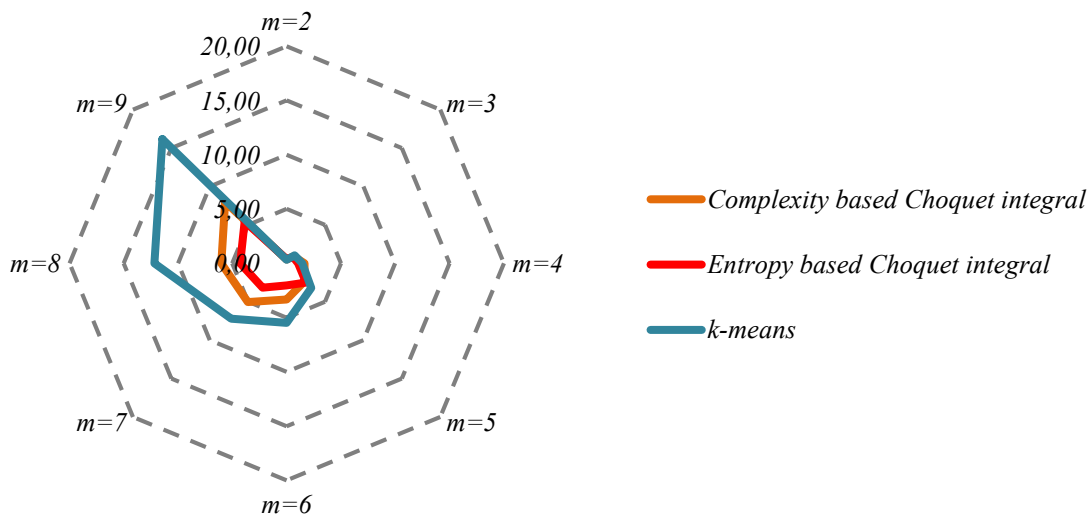


Figure 3. The *MSE* results of the methods

The *MSE* results of k -means method are respectively 0.33, 1.03, 1.45, 3.21, 5.48, 7.21, 12.18, and 16.19 for $c = 2, 3, 4, 5, 6, 7, 8$, and 9. k -means results can be compared with only “equal threshold” approach results. As the number of the cluster increases, it is seen that the *MSE* value increases. Besides, it can be seen that as the number of the level of score increases, *MSE* value increases. Nevertheless, if “equal threshold” method is used, this increase is less than it is if k -means method is used. By using the idea of the model with the smallest *MSE* value, the results of the robustness analysis indicate that both entropy and complexity based discrete Choquet integral provides better results than k -means method in most of the cases as shown in Figure 3.

4. CONCLUSIONS and REMARKS

Evaluation of the academic performance, that takes a wide variety of methods, is an integral part of educational system. That evaluation depends on many criteria that can be seen as a MCDM problem. These problem refers to the analysis and judgment process of selecting an optimal solution from two or more feasible schemes with multiple indicators in order to achieve a certain goal. As for the Choquet integral operator of fuzzy measure, since Schmeidler (1989) first applied it to related MCDM analysis, it has been widely used in decision-making fields for performance evaluation such as engineering, economy and management areas (Xu, 2010; Sun et al., 2015; Han & Wei, 2017; Liu et al., 2018).

At the present time, most of the traditional evaluation techniques take no account of the interactions among criteria. In this regard, the Choquet integral is an effective and appropriate method drawing strong attention to inherently dependent evaluation criteria. In this study, an extensive comparison of several discretization techniques is mapped out for objectively evaluating academic performance of the students. In detail, the discrete Choquet integral is used with the ultimate aim of evaluating the students' success at a university in Turkey. Even though, a specific framework is provided, the method can also be used in any educational assessment such as teacher competency in higher institution of learning and universities perform according to different educational indicators. Thus, the method can be seen as a tool that attracts a good deal of attention in educational assessment.

In this study, the entropy-based and the complexity-based discrete Choquet integral and the k -means method is used. For the ex-post evaluation, the mean square error method is used in our study. Previous works on the evaluation of students' performance by using the discrete Choquet integral such as Shieh et al., and Chang et al., (2009) did not consider whether the data matrix was normally distributed. However, this study showed that if the data matrix is not normally distributed, entropy-based Choquet integral provides much better results. On the other hand, complexity-based Choquet integral generally presents optimal results if the data is close to being normally distributed. Besides, the other previous studies can show a good performance and a good accuracy results when the sample size is large, but it cannot be possible to deal with the problems when the size is small. Another important aspect of our evaluation is that the paper presents the k -means method as a robustness analysis to compare the effectiveness of the discrete Choquet integral based methods. The most remarkable property of k -means is its efficiency in large sample size. However, the obtained mean square error results of the k -means method indicate that both entropy and complexity based Choquet integral method provides better results than the k -means method in most of the cases. In conclusion, this study's findings point out that the discrete Choquet integral method provides a major support to educational system in evaluating students' performance.

ORCID

Deniz Koçak  <https://orcid.org/0000-0002-5893-0564>

5. REFERENCES

- Angilella, S., Arcidiacono, S.G., Corrente, S., Greco, S., & Matarazzo, B. (2017). An application of the SMAA–Choquet method to evaluate the performance of sailboats in offshore regattas. *Operational Research*, doi:10.1007/s12351-017-0340-7
- Angilella, S., Corrente, S., & Greco, S. (2015). Stochastic multiobjective acceptability analysis for the Choquet integral preference model and the scale construction problem. *European Journal of Operational Research*, 240(1), 172 - 182, doi: 10.1016/j.ejor.2014.06.031
- Baker, R.S.J.D., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, 1(1), 3-17.

- Branke, J., Corrente, S., Greco, S., Słowiński, R., & Zielniewicz, P. (2016). Using Choquet integral as preference model in interactive evolutionary multiobjective optimization. *European Journal of Operational Research*, 250(3), 884-901, doi: [10.1016/j.ejor.2015.10.027](https://doi.org/10.1016/j.ejor.2015.10.027)
- Calvo, T., Mayor, G., & Mesiar, R. (2002). *Aggregation operators: New trends and applications*. Physica-Verlag, Heidelberg, New York.
- Catlett, J. (1991). *On changing continuous attributes into ordered discrete attributes*. Kodratoff, Y. (Eds.) Machine Learning Lecture Notes in Computer Science, Springer, Berlin, Heidelberg.
- Chang, H.J., Liu, H.C., Tseng, S.W., & Chang, F.M. (2009). A comparison on Choquet integral with different information-based fuzzy measures. In *Proceedings of the 8th International Conference on Machine Learning and Cybernetics* (pp. 3161-3166).
- Chmielewski, M.R., & Grzymala-Busse, J.W. (1996). Global discretization of continuous attributes as preprocessing for machine learning. *International Journal of Approximate Reasoning*, 15(4), 319-331. doi: [10.1016/S0888-613X\(96\)00074-6](https://doi.org/10.1016/S0888-613X(96)00074-6)
- Choquet, G. (1954). Theory of capacities. *Annals of Institute of Fourier*, 5, 131-295.
- Cui, L., & Li, Y. (2008). Linguistic quantifiers based on Choquet integrals. *International Journal of Approximate Reasoning*, 48(2), 559-582. doi: [10.1016/j.ijar.2007.11.001](https://doi.org/10.1016/j.ijar.2007.11.001)
- Dougherty, J., Kohavi, R., & Sahami, M. (1995). Supervised and unsupervised discretization of continuous features. In *12th International Conference on Machine Learning*. Los Altos, CA: Morgan Kaufmann, (pp. 194-202).
- Flynt, A., & Dean, N. (2016) A survey of popular R packages for cluster analysis. *Journal of Educational and Behavioral Statistics*, 41(2), 205-225. doi: [10.3102/1076998616631743](https://doi.org/10.3102/1076998616631743)
- Garcia, S., Luengo, J., Saez, A., Lopez, V., & Herrera, F. (2013). A survey of discretization techniques: Taxonomy and empirical analysis in supervised learning. *IEEE Transactions on Knowledge and Data Engineering*, 25(4), 734-750. doi: [10.1109/TKDE.2012.35](https://doi.org/10.1109/TKDE.2012.35)
- Grabisch, M. (1996). The application of fuzzy integrals in multicriteria decision making. *European Journal of Operational Research*, 89(3), 445-456, doi: [10.1016/0377-2217\(95\)00176-X](https://doi.org/10.1016/0377-2217(95)00176-X)
- Greene, W.H. (2016). *Econometric Analysis*. 7th Edition, Pearson Education, Inc., Publishing as Prentice Hall. NJ 074458, USA.
- Han, L., & Wei, C. (2017). Group decision making method based on single valued neutrosophic Choquet integral operator. *Operations Research Transactions*, 21(2), doi: [10.15960/j.cnki.issn.1007-6093.2017.02.012](https://doi.org/10.15960/j.cnki.issn.1007-6093.2017.02.012)
- Herde, C.N., Wüstenberg, S., & Greiff, S. (2016). Assessment of complex problem solving: What we know and what we don't know. *Applied Measurement in Education*, 29(4), 265-277, doi: [10.1080/08957347.2016.1209208](https://doi.org/10.1080/08957347.2016.1209208)
- Huber, S.G., & Skedsmo, G. (2016). Teacher evaluation accountability and improving teaching practices. *Educational Assessment, Evaluation and Accountability*, 28(3), 105-109, doi: [10.1007/s11092-016-9241-1](https://doi.org/10.1007/s11092-016-9241-1)
- Jain, A.K. (2010). Data clustering: 50 years beyond k-means. *Pattern Recognition Letters*, 31(8), 651-666, doi: [10.1016/j.patrec.2009.09.011](https://doi.org/10.1016/j.patrec.2009.09.011)
- Kasparian, J., & Rolland, A. (2012). OECD's 'Better Life Index': Can any country be well ranked?. *Journal of Applied Statistics*, 39(10), 2223 - 2230, doi: [10.1080/02664763.2012.706265](https://doi.org/10.1080/02664763.2012.706265)
- Kononenko, I., & Kukar, M. (2007). *Machine learning and data mining: Introduction to principles and algorithms*. Harwood Publishing Limited.
- Liu, W.F., Du Y. X., & Chang J. (2018). Intuitionistic fuzzy interaction choquet integrals operators and applications in decision making. *Fuzzy systems and Mathematics*, 32(2), doi: [1001-7402\(2018\)02-0110-11](https://doi.org/10.1001-7402(2018)02-0110-11)

- Liu, H., Hussain, F., Tan, C.L., & Dash, M. (2002). Discretization: An enabling technique. *Data Mining and Knowledge Discovery*, 6(4), 393-423, doi: [10.1023/A:1016304305535](https://doi.org/10.1023/A:1016304305535)
- Marichal, J.L., & Roubens, M. (2000). Determination of weights of interacting criteria from a reference set. *European Journal of Operational Research*, 124(3), 641-650, doi: [10.1016/S0377-2217\(99\)00182-4](https://doi.org/10.1016/S0377-2217(99)00182-4)
- Peña-Ayala, A. (2014). Educational data mining: A survey and a data mining-based analysis of recent works. *Expert Systems with Applications*, 41(4), 1432-1462, doi: [10.1016/j.eswa.2013.08.042](https://doi.org/10.1016/j.eswa.2013.08.042)
- Pötzelberger, K., & Felsenstein, K. (1993). On the fisher information of discretized data. *The Journal of Statistical Computation and Simulation*, 46(3-4), 125-144.
- Pyle, D. (1999). *Data Preparation for Data Mining*. Morgan Kaufmann Publishers, Inc.
- Schmeidler D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, 57(5), 571-587, doi: [10.2307/1911053](https://doi.org/10.2307/1911053)
- Shieh, J.I., Wu, H.H., & Liu, H.C. (2009). Applying a complexity-based Choquet integral to evaluate students' performance. *Expert Systems with Applications*, 36(3), 5100-5106, doi: [10.1016/j.eswa.2008.06.003](https://doi.org/10.1016/j.eswa.2008.06.003)
- Slater, S., Joksimovic, S., Kovanovic, V., & Baker, R.S. (2017). Tools for educational data mining: A review. *Journal of Educational and Behavioral Statistics*, 42(1), 85-106, doi: [10.1016/j.eswa.2008.06.003](https://doi.org/10.1016/j.eswa.2008.06.003)
- Sun, H. X., Yang, H. X., Wu, J. Z., & Ouyang, Y. (2015). Interval neutrosophic numbers Choquet integral operator for multi-criteria decision making. *Journal of Intelligent & Fuzzy Systems*, 28(6), 2443-2455.
- Tan, P.N., Steinbach, M., & Kumar, V. (2005). *Introduction to data mining*. Pearson, Addison Wesley.
- Veeramuthu, P., Periyasamy, R., & Sugasini, V. (2014). Analysis of student result using clustering techniques *International Journal of Computer Science and Information Technologies*, 5(4), 5092-5094.
- Wang, R.-S. & Ha, M.-H. (2008). On the properties of sequences of fuzzy-valued Choquet integrable functions. *Fuzzy Optimization and Decision Making*, 7, 417-431, doi: [10.1007/s10700-008-9040-3](https://doi.org/10.1007/s10700-008-9040-3)
- Wang, Z., Nian, Y., Chu, J., & Shi, Y. (2012). *Emerging Computation and Information Technologies for Education*. Mao, E., Xu, L., & Tian, W. (Eds.), Springer Verlag, Berlin Heidelberg.
- Wen, X., Yan, M., Xian, J., Yue, R. & Peng, A. (2016). Supplier selection in supplier chain management using Choquet integral-based linguistic operators under fuzzy heterogeneous environment. *Fuzzy Optimization and Decision Making*, 15, 307-330.
- Xu, Z. (2010). Choquet integrals of weighted intuitionistic fuzzy information. *Information Sciences*, 180(1), 726-736, doi: [10.1016/j.ins.2009.11.011](https://doi.org/10.1016/j.ins.2009.11.011)
- Zopounidis, C., Doumpos, M. (2002). Multicriteria classification and sorting methods: A literature review. *European Journal of Operational Research*, 138(2), 229-246, doi: [10.1016/S0377-2217\(01\)00243-0](https://doi.org/10.1016/S0377-2217(01)00243-0)