IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 17 October 2023, accepted 28 October 2023, date of publication 6 November 2023, date of current version 10 November 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3330245

RESEARCH ARTICLE

Quadrature-Inspired Generalized Choquet Integral in an Application to Classification Problems

PAWEŁ KARCZMAREK[1], MICHAŁ DOLECKI[1], PAWEŁ POWROZNIK[2], (Member, IEEE), ZBIGNIEW A. ŁAGODOWSKI[3], ADAM GREGOSIEWICZ[3], ŁUKASZ GAŁKA[1], WITOLD PEDRYCZ[4,5], (Life Fellow, IEEE), DARIUSZ CZERWINSKI[6], (Member, IEEE), AND KAMIL JONAK[7]

[1]Department of Computational Intelligence, Lublin University of Technology, 20-618 Lublin, Poland
[2]Department of Computer Science, Lublin University of Technology, 20-618 Lublin, Poland
[3]Department of Mathematics, Lublin University of Technology, 20-618 Lublin, Poland
[4]Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6R 2V4, Canada
[5]Systems Research Institute, Polish Academy of Sciences, 01-224 Warsaw, Poland
[6]Department of Applied Computer Science, Lublin University of Technology, 20-618 Lublin, Poland
[7]Department of Clinical Neuropsychiatry, Medical University of Lublin, 20-059 Lublin, Poland

Corresponding author: Paweł Karczmarek (p.karczmarek@pollub.pl)

This work was supported by the Internal Fund of the Scientific Discipline under Grant FD-20/IT-3/004 and Grant FD-20/IT-3/047.

ABSTRACT Correct classification remains a challenge for researchers and practitioners developing algorithms. Even a minor enhancement in classification quality, for instance, can significantly boost the effectiveness of detecting conditions or anomalies in safety data. One solution to this challenge involves aggregating classification results. This process can be executed effectively as long as the aggregation function is appropriately chosen. One of the most efficient aggregation operators is the Choquet integral. Furthermore, there exist numerous generalizations and extensions of the Choquet integral in the existing literature. In this study, we conduct a comprehensive analysis and evaluation of a novel approach for deriving an aggregate classification. The aggregation process applied to various classifiers is based on enhancements to the Choquet integral. These novel expressions draw inspiration from Newton-Cotes quadratures and other well-known formulae from numerical analysis. In contrast to previous approaches that exploit the generalization of the Choquet integral, our approach requires the utilization of two or three adjacent values associated with the membership of a specific element in different classes. This enables the use of more efficient enhancements in terms of accuracy measurement. Specifically, the t-norm following the integral symbol can be effectively replaced by mathematical expressions used in executing numerical integration formulae. This leads to more precise results and aligns with the concept of numerical integration. Furthermore, in a series of experiments, we thoroughly assess the performance of the proposed approach in terms of classification accuracy. We analyze the strengths and weaknesses of the new approach and establish the experimental settings that can be applied to similar tasks. In the series of experiments, we have demonstrated that the proposed Quadrature-Inspired Generalized Choquet Integral (QIGCI) can either outperform previous enhancements of the Choquet integral or at least achieve a similar level of accuracy measurement. However, we also highlight scenarios where previous approaches can still be a suitable choice. The number of QIGCI-based aggregation models that outperform others is convincing, indicating that this approach is worthy of consideration.

INDEX TERMS Aggregation, anomaly detection, Choquet integral, generalized Choquet integral, numerical quadratures, many classifiers, pre-aggregation functions.

The associate editor coordinating the review of this manuscript and approving it for publication was Nazar Zaki.

I. INTRODUCTION

The The problem of classification has recently become one of the most significant challenges in theoretical and applied computer science. It arises in various contexts such as image analysis, anomaly detection, and tabular data recognition. Especially in the past decade, this task has seen increasingly effective solutions thanks to machine learning algorithms based on deep learning. Classical methods like decision trees are also frequently employed and have been successfully enhanced. The field is multifaceted, with ongoing advancements including the development of more efficient processors and graphics cards, algorithm refinements, and code optimization. However, we still lack ideal methods. One potential candidate for a universal solution is the aggregation of classification results from various classifiers. Proper utilization of aggregation operators has the potential to greatly enhance method efficiency. These operators depend on several factors, with the most critical ones being the weights assigned to individual classifiers (which can be determined through expert assessment, pre-tests, or results from similar problems) and the measurement of the degree of an element's membership in a class. This measurement can be implemented as a function of probability, distance, fuzzy membership, ranking, voting, and other criteria.

There is a wide range of aggregation methods available. Common operators encompass various types of averages, the median, voting, maximum, and minimum. However, numerous other, more sophisticated methods are comprehensively detailed in monographs [1], [2], [3], [4], [5], [6], articles [7], [8], [9], [10], or reviews, e.g. [11]. Special classes of aggregation functions are distinguished in the literature. These include t-norms [12], [13], OWA operators [14], [15], fuzzy measures [16], polynomial function [17], order-2 fuzzy sets [18], or fuzzy generalized unified aggregation operator [19]. There are also granular models, see [20] and [21]. Finally, Choquet integral (CI) [22], [23] and so-called pre-aggregation operators have been recently widely discussed [24], [25], [26], [27], [28], [29], [30] including their interval-valued versions [31]. Obviously, there are also proposed various other similarly built operators, e.g. Choquet-like copula [32], recent generalizations of Choquet integral mixed with overlap functions [33], or neutrosophic Choquet integral [34], Fermatean hesitant fuzzy Choquet operator [35], or Shapley-based generalizations of CI [37], bidirectional CI [38], or CI with a spherical fuzzy set [39]. Generally speaking, the whole pre-aggregation function class are functions based on replacing the multiplication operator by the t-norm under the integral symbol. An in-depth experimental comparison of the first results obtained for this class of functions was presented, among others, in [36] and [40]. In general, the problem of selection of optimal aggregation function has been also considered recently in [41], [42], [43], and [44].

Typically, generalizations and extensions of the Choquet integral, which are utilized as aggregation functions, yield better results than the Choquet integral itself, often regarded as one of the finest aggregators. A comprehensive overview of this topic can be found in [28] and other references. On one hand, it is important to note that the results may be slightly affected by factors like data distribution, classifier selection, or their respective weights. However, it can be argued that replacing the algebraic product with one of the frequently used t-norms, such as the Łukasiewicz norm, Hamacher norm, minimum, maximum, and others, should enhance the quality of classification when relying on multiple classifiers.

Hence, this study main purpose is to propose a novel approach to the problem of classification realized by the Choquet integral on a basis of the quadrature formulae well-known from the numerical analysis theory. Using such kind of formulae, one can obtain the Choquet-like integral more precisely or accurately. The next goal is to propose the best set of operator combinations appearing in anomaly detection or multi-class problems. Such a combination may depend on a choice of t-norms, integration rule, or t-norm parameter value. It is worth stressing that this study is a significantly extended version of the conference paper [45] which was presented in FUZZ-IEEE Conference. In the conference paper, mainly the databases used to detect anomalies were discussed. Here, the study is augmented in several ways. First, we provide the mathematical properties of the Quadrature-Inspired Generalizations of Choquet Integral (QIGCI). Next, the experimental results comprehensively show the performance of the aggregation operators in dependence on various families of t-norms and their parameters put under the integral symbol. Particular attention is paid to popular and often discussed t-norms such as algebraic or drastic product, Łukasiewicz, Hamacher, minimum, and nilpotent minimum functions. Finally, we also consider 27 additional databases used to verify multi-class classification algorithms to fully examine and describe the potential of the proposed approach since in the previous publication only two-class task was realized. Moreover, we report the results experiments carried multiple number of times to find the standard deviations and demonstrate the proposed aggregation formulae stability.

In summary, the main new points presented in this paper include an introduction of Quadrature-Inspired Generalization of Choquet Integral as well as its application to the problem of multi-class classification. Moreover, we prove in a series of experiments that the new approach is better than the previous ones enhancements of Choquet integral in terms of accuracy measure.

The structure of the study is as follows. In Section II, the CI theory and its remarkable generalizations are recalled. The new enhancements of the Choquet integral are discussed in Section III. Section IV is devoted to experimental results. Conclusions and possible directions of future work are detailed in Section V. Finally, we include detailed formulae (Appendix).

II. CHOQUET INTEGRAL AND ITS ENHANCEMENTS

Let us recall the main properties of the CI. Here, X is a set. Then $Q(X) = 2^X$ is a family of all X subsets. Let $m : Q(X) \to \mathbb{R}$ be a function. Under the following conditions the function m is called a fuzzy measure.

$$m(\emptyset) = 0 \tag{1}$$
$$m(X) = 1 \tag{2}$$
$$m(A) \leq m(B), \quad A \subset B, \quad A, B \in Q(X) \tag{3}$$
$$[-2pt] \lim_{n\to\infty} m(A_n) = m\left(\lim_{n\to\infty} A_n\right) \tag{4}$$

where $\{A_n\}$, $n = 1, 2, \ldots$ is an increasing set sequence which means that $A_1 \subset A_2 \subset \ldots$. However, the property (4) is not required for a finite number of classifiers as in this text.

Next, the Sugeno fuzzy measure, also called λ-fuzzy measure, satisfies the equations

$$m(A \cup B) = m(A) + m(B) + \lambda m(A) m(B) \tag{5}$$

where $\lambda > -1$. Note that A and B do not non-overlap. Furthermore,

$$m(A_{i+1}) = m(A_i) + m_{i+1} + \lambda m(A_i) m_{i+1} \tag{6}$$

and $A_i = \{t_1, \ldots, t_i\}$, $A_{i+1} = \{t_1, \ldots, t_{i+1}\}$. Note that hereafter we use the common simplified form of this notation, namely

$$m_i = m(\{t_i\}), \quad i = 1, \ldots, n \tag{7}$$

It is worth noting that a subset can be identified with a classifier and the measure defined for this subset is a measure of the significance of a given classifier. All classifiers, combined in some way, should intuitively have the highest measure of importance if $\lambda > 0$. The considerations of this problem were given in [46].

Assume that $h(t) : Q(X) \to [0, 1]$ is a function which is non-increasing. Namely, $h(t_{i+1}) \geq h(t_i)$, $i = 1, \ldots, n$. Also, one has to assume $h(t_{n+1}) = 0$. Then the CI in its generic for is given by

$$Ch = \sum_{i=1}^{n} \left(\left(h(t_i) - h(t_{i+1})\right) m(A_i)\right) \tag{8}$$

Along with (8) we encounter a number of extensions. Let us recall a few of them.

In a series of works [24], [25], [47], [48], and [49] introduced were various generalization kinds of (8). All of them are often called pre-aggregation functions. The main formulae are expressed as

$$Ch_T = \sum_{i=1}^{n} \left(T\left(h(t_i) - h(t_{i+1})\right), m(A_i)\right) \tag{9}$$

$$Ch_F = \min\left(\sum_{i=1}^{n} T\left(h(t_i) - h(t_{i+1}), m(A_i)\right), 1\right) \tag{10}$$

In [48] a Choquet-like operator

$$Ch_T C = \sum_{i=1}^{n} \left(T\left(h(t_i), m(A_i)\right) - T\left(h(t_{i+1}), m(A_i)\right)\right) \tag{11}$$

was proposed. In [50] a t-norm T was substituted by min operator. Here, we denote it C_{min}. In [51] an overlapping operator O_v was put in the place of T. O_v is commutative, increasing, and continuous. It fulfills the conditions $O_v(t, s) = 0$ for $ts = 0$ and $O_v(t, s) = 1$ for $ts = 1$. In [52] and in the monograph [40] the following generalizations were proposed:

$$Ch_{TTC} = \sum_{i=1}^{n} \left(T\left(h(t_i), m(A_i)\right) - T\left(h(t_{i+1}), m(A_i)\right) + T\left(h(t_i) - h(t_{i+1}), m(A_i)\right)\right) \tag{12}$$

$$Ch_{Tmin} = \sum_{i=1}^{n} T\left(\min\left(h(t_i), m(A_i)\right) - \min\left(h(t_{i+1}), m(A_i)\right), m(A_i)\right) \tag{13}$$

$$Ch_{minT} = \sum_{i=1}^{n} \min\left(T\left(\left(h(t_i), m(A_i)\right)\right), T\left(h(t_{i+1}), m(A_i)\right)\right) \tag{14}$$

$$Ch_{Diff_1} = \sum_{i=1}^{n} T\left(h(t_{i-1}) - h(t_{i+1}), m(A_i)\right) \tag{15}$$

$$Ch_{Diff_2} = \sum_{i=1}^{n} T\left(h(t_{i-1}) + h(t_{i+1}) - h(t_i), m(A_i)\right) \tag{16}$$

and

$$Ch_{Diff_3} = \sum_{i=1}^{n} T\left(\left(h(t_{i-1}) - h(t_{i+1})\right)/h(t_i), m(A_i)\right) \tag{17}$$

III. CHOQUET INTEGRALS INSPIRED BY QUADRATURES

The concept of Choquet integral is relatively easy to generalize and extend. Moreover, to carry out the generalization operations, the computer program does not need new data, but only minor modifications to the formulae themselves. Let us take a closer look at this concept. Particularly, one can observe that in formula (8), the successive differences in the $h(\cdot)$ function values are most important. They are sorted in a non-increasing way in the arguments t_i. This form of the formula encourages actions where one can attempt to specify changes between the adjacent function $h(\cdot)$ values. In addition, it would be interesting to utilize all values of $h(\cdot)$ to build the parameter of $T(\cdot)$. This conclusion as well as possible changes of the Ch_{Diff_i}, $i \in \{1, 2, 3\}$, form may lead to the statement that one can apply much wider range of the function $h(t_i)$ values as this parameter. In a more plastic way, it can be said that using more points to evaluate this expression can lead to its more precise value, as well as reflect the idea of numerical calculations for various integral forms.

TABLE 1. Details of the multi-class problems databases.

Dataset/No.	Classes	Records
Covtype	7	116202
Skin Nonskin	2	49011
Shuttle	7	11600
Letter Recognition	26	4000
Isolet	26	1559
Online News Popularity	2	7929
Dry Bean	7	2722
Drift	6	2782
Abalone	28	835
EEG Eye State	2	2996
Satimage	6	1287
Wine Quality White	7	980
Page Blocks	5	1095
UCI Named Data	2	2000
Predict Students	3	885
Segment	7	462
Spambase	2	920
Wine Quality Red	6	320
Onehr	2	370
Eighthr	2	369
Erman Numeric	2	200
Breast Cancer Wisconcin	2	137
Glass	6	43
Ionosphere	2	70
Zoo	7	20
Wine	3	36
Iris	3	30
Lung Cancer	3	5

A helpful function class can be here quadratures [53] or formulae coming from various disciplines such as Program Analysis Review Technique (PERT) Formula of project time or cost estimation. We focus here on a class of functions called Newton-Cotes (N-C) quadratures. Mathematicians use them to determine the value of the definite integral defined on an interval. The calculation is realized using equal segments of the interval. The quadrature formulae are mainly divided into two classes of operators, namely closed formulae when the boundary values of an interval are utilized or open in an opposite case.

Now, we take a more closer look at the solution we propose. Consider a function

$$Ch_Q = \sum_{i=1}^{n} \left(T\left(Q\left(h_i\right), m\left(A_i\right)\right)\right) \tag{18}$$

The equation (18) can be used as a general framework to build Choquet integral enhancements. To be precise, an expression $Q(h_i)$ can be understood as follows. If Q is related to a Simpson rule or the above-mentioned PERT equation then

$$Q_S\left(h_i\right) = \frac{1}{6}\left(h_{i-1} + 4h_i + h_{i+1}\right) \tag{19}$$

Here, $h_i = h(t_i)$. Hence, the generalized CI (18) is

$$Ch_{Q_S}(t) = \sum_{i=1}^{n} \left(T\left(\frac{1}{6}\left(h_{i-1}(t) + 4h_i(t) + h_{i+1}(t)\right), m\left(A_i\right)\right)\right) \tag{20}$$

The Simpson rule is obviously slightly modified here. The weights are normalized so that their sum is 1. The list of other rules is presented in the Appendix. Moreover, to retain the monotonicity property of the function $h(\cdot)$ we assume that the values h_i for $i < 0$ are equal to h_0. Similarly, $h_i = 0$ for $i \geq n+1$.

TABLE 2. Top 5 best average percentage values over datasets wrt. kind of task.

Gen. of CI	T-norm	α	Accuracy
Anomaly detection			
N-C 11-point	7	3.5	95.53
Open N-C 6-point	11	1.1	95.52
Hardy	25	2.4	95.52
CDiff1	14	1.3	95.51
N-C 7-point	25	2.4	95.50
Multi-class problem			
Extrapolative N-C 2-point	23	6.3	92.07
CDiff2	3	-8.9, -9.9	92.03
Open N-C 2-point	23	6	92.02
Ov			91.96
CT, CF	11, 14	[9, 10]	91.94
CTMin, CMinT	11, 14	[8.2, 10]	91.94

It is worth noting that despite many comparisons of various quadratures [54], [55] for numerical integration problem and their known error estimates, they are analyzed experimentally, in terms of applications to slightly different kind of problem. We do not choose only a few of them which could be suspected to be more effective than others. Finally, it is also worth stressing that the CI enhancements can be understood as some expansions of the integral of the Sugeno form since after the integral symbol there is only a single quadrature and no difference of quadratures appears.

IV. MATHEMATICAL PROPERTIES OF NEW CI GENERALIZATIONS

We investigate properties of the function Ch_Q given by (18). In particular, refer to Theorem 1, we prove that Ch_Q is an $\vec{r}$-non-decreasing function under mild conditions on T and Q. Here, given a vector $\vec{r} \in [0,1]^n$, by an $\vec{r}$-non-decreasing function, as defined in [24], we consider a function $F\colon [0,1]^n \to [0,1]$ acting on the n-th dimensional cube $[0,1]^n$ with values in the unit interval $[0,1]$ such that $F(\vec{t}) \leq F(\vec{t} + c\vec{r})$ for all $\vec{t} \in [0,1]^n$ and all $c > 0$ satisfying $\vec{t} + c\vec{r} \in [0,1]^n$.

First we establish a technical lemma. We consider $[0,1]^n$ as a space equipped with the coordinatewise order, that is for $\vec{t} = (t_1, \ldots, t_n) \in [0,1]^n$ and $\vec{s} = (s_1, \ldots, s_n)$ we say that $\vec{t} \leq \vec{s}$ if and only if $t_i \leq s_i$ for all $i \in \{1, \ldots, n\}$. In particular, the fact that a function $F\colon [0,1]^n \to [0,1]$ is non-decreasing means that $F(\vec{t}) \leq F(\vec{s})$ for $\vec{t}$ and $\vec{s}$ as above.

Lemma 1: Let $T\colon [0,1] \to [0,1]$ be a non-decreasing function with respect to the first coordinate: for all $t, t' \in [0,1]$ such that $t \leq t'$ it follows that $T(t,s) \leq T(t',s)$ for all $s \in [0,1]$. Moreover, for every $i \in \{1, \ldots, n\}$ let $Q_i\colon [0,1]^n \to [0,1]$ be a non-decreasing function. Then for every $s \in [0,1]$ the function $F\colon [0,1]^n \to [0,1]$ defined by

$$F(\vec{t}) = \sum_{i=1}^{n} T(Q_i(\vec{t}), s)$$

is non-decreasing.

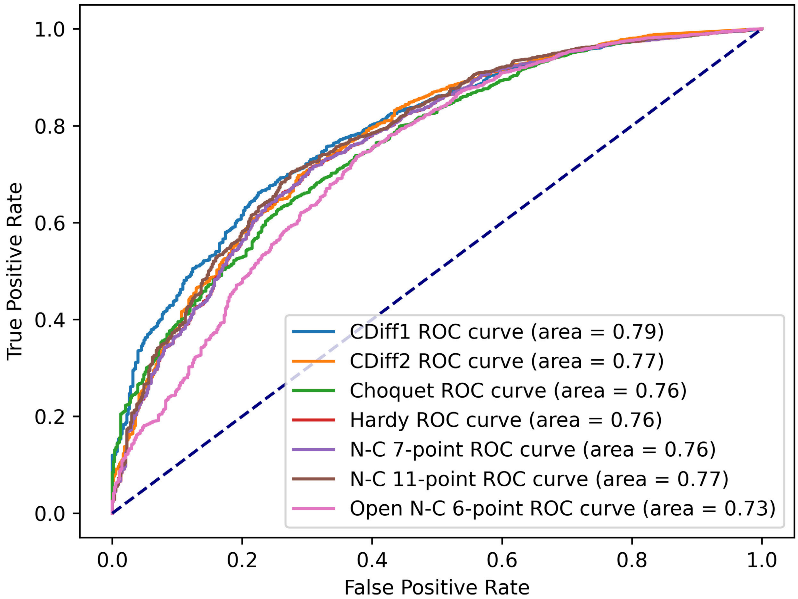


FIGURE 1. ROC curves and AUC measures obtained with the best aggregation operators.

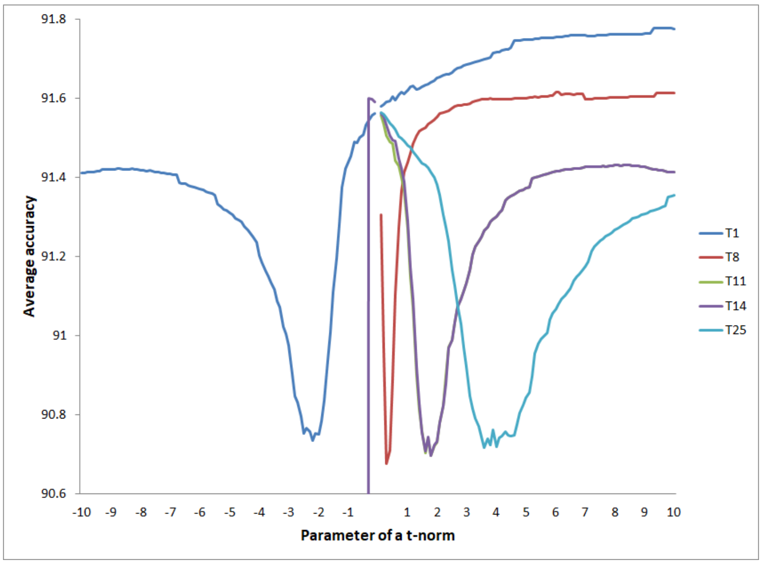


FIGURE 2. Selected t-norms and their average accuracy performance in the multi-class problems when the Open N-C 2-point quadrature is used.

Proof: The conclusion follows by the fact that F is a sum of compositions of non-decreasing functions.

As a simple consequence of Lemma 1 we obtain the following corollary.

Corollary 1: Let T be as in Lemma 1 and suppose that Q is an integral quadrature appearing in (18) with positive coefficients. Then Ch_Q given by (18) is an $\vec{r}$-non-decreasing function for all $\vec{r} \in [0, 1]^n$. Moreover, $Ch_Q = 0$ for $\vec{h} = \vec{0}$, provided that $T(0, s) = 0$ for all $s \in [0, 1]$.

Proof: Because Q has positive coefficients, the function $\vec{h} \mapsto Q(h_i)$ is non-decreasing for all $i \in \{1, \ldots, n\}$. Moreover, for $\vec{h} = \vec{0}$ we have $Ch_Q = \sum_{i=1}^{n} T(0, m(A_i)) = 0$.

By applying Corollary 1 we prove that Ch_Q is an $\vec{r}$-non-decreasing function for a big function class T, namely t-norms. Here, by a t-norm, as in [1], we consider a function $T : [0, 1]^2 \to [0, 1]$ satisfying the conditions:

1) $T(t, s) = T(s, t)$ for all $t, s \in [0, 1]$,
2) $T(T(t, s), z) = T(t, T(s, q))$ for all $t, s, q \in [0, 1]$,
3) $T(t, s) \leq T(t', s')$ for all $t, t', s, s' \in [0, 1]$ satisfying $t \leq t'$ and $s \leq s'$,
4) $T(t, 1) = t$ for all $t \in [0, 1]$.

Commonly encountered examples of T-norms include (here, naturally, $t, s \in [0, 1]$, see [24]):

TABLE 3. The best average percentage accuracy values wrt. each dataset. Anomaly detection task.

Dataset	CI generalization	T-norm no.	α	Acc.	Std.
Annthyroid	CDiff2	23	6.6, 6.7	96.60	0.11
Arrhythmia	Ov			98.76	0.18
Cardio	Open N-C 2-point	1	-10	97.23	0.03
Glass	Extr. N-C 3-point	10	[4.8, 5.1]	99.30	0.31
Ionosphere	CDiff2	23	3	92.82	0.40
Letter	CF	4	0.1	96.35	0.20
Lympho	Hardy, Weddle, N-C 7-point	23	6.5	97.77	0.43
Mammography	Extr. N-C 2-point	10	9.8	97.17	0,10
Musk	Simpson	23	9.4	96.98	0.26
Optdigits	CDiff2	23	5.2	98.37	0.22
Pendigits	CTMin, CMinT	23	5.2	98.93	0.08
Pima	Hardy	24	0.3	95.98	0.19
Satellite	CT	23	4.9	97.72	0.20
Satimage-2	N-C 10-point	23	10	96.77	0.07
Thyroid	Ov			98.05	0.13
Vertebral	Extr. N-C 3-point	1	-4	96.92	0.50
Vowels	N-C 9-point	14	-9	93.47	1.13
Wbc	Open N-C 6-point	7	5.4, 5.5	100.00	0.00
Wine	Simpson	23	*	98.45	0.00
	Bool				
	N-C 3-point				
	Newton-Cotes 4				
	Hardy				
	Durand				
	Shovelton				
	Weddle				
	Woolhouse				
	N-C 6-point				
	N-C 10-point				
	Open N-C 3-point				
	Extr. N-C 2-point				
	Extr. N-C 3-point				
	Open N-C 7-point	5, 20	*		
	Ov				

1) Minimum

$$T_M(t, s) = \min\{t, s\} \tag{21}$$

2) Algebraic product

$$T_P(t, s) = ts \tag{22}$$

3) Łukasiewicz t-norm

$$T_L(t, s) = \max\{t + s - 1, 0\} \tag{23}$$

4) Drastic product

$$T_{DP}(t, s) = \begin{cases} t & \text{for } s = 1 \\ s & t = 1 \\ 0 & \text{otherwise} \end{cases} \tag{24}$$

5) Nilpotent minimum

$$T_{NM}(t, s) = \begin{cases} \min\{t, s\} & \text{for } t + s > 1 \\ 0, & \text{otherwise} \end{cases} \tag{25}$$

6) Hamacher product

$$T_{HP}(t, s) = \begin{cases} 0, & \text{for } t = s = 1 \\ \dfrac{ts}{t + s - ts} & \text{otherwise} \end{cases} \tag{26}$$

and all t-norms T_i for $i \in \{1, \ldots, 25\}$ defined in Section 2.6 in [1].

Finally, we state the concluding result, which is a consequence of Corollary 1 and the definition of a t-norm.

Theorem 1: Let T be a t-norm and suppose that Q is an integral quadrature appearing in (18) with positive coefficients. Then Ch_Q is $\vec{r}$-non-decreasing for all $\vec{r} \in [0, 1]^n$, and moreover, $Ch_Q = 0$ for $\vec{h} = \vec{0}$.

It is worth noting that from Corollary 1 it follows that the class of overlap functions [56], apart from t-norms, can also be considered here, see previous section for the details of this class.

V. NUMERICAL EXPERIMENTS

Let us consider the results of experiments carried with different enhancements of Choquet integral. We are interested in finding the best aggregation operators built using Quadrature-Inspired Generalizations of Choquet Integral, the t-norms with which they are composed, as well as the parameters of those t-norms for which the best average results were obtained. We are also interested in the best results for individual bases related to anomaly detection and multi-class problems. Finally, we analyze in detail the results for the t-norms of the form (21)-(26) commonly described in the literature.

A. DATASETS USED IN THE EXPERIMENTS

Here, we briefly recall the datasets utilized in our experimental series. ODDS is one of the most often used set of data tables applied to analyze anomaly detection algorithms.

TABLE 4. The best average percentage accuracy values wrt. each dataset. Multi-class problem.

Dataset	CI generalization	T-norm no.	α	Acc.	Std.
Abalone	CT, CF	1, 8, 11, 14	*	93.26	0.87
	CTMin, CMinT, CDiff2	1, 8, 11, 14, 25			
Breast C.	Ov			94.53	0.75
Data UCI	Ov			94.53	0.35
Drift	CT, CF, CTMin, CMinT, CDiff2	1, 8, 11, 14, 25	*	93.13	1.05
	CDiff1	11, 14			
Dry B.	CT, CF, CTMin, CMinT	1, 8, 11, 14	*	93.30	1.07
	CDiff2	1, 8, 11, 14, 25			
	Open N-C 2-point	11			
EEG	Ov			93.36	0.54
Eighthr	Extr. N-C 4-point	8	0.3	92.47	1.33
Erman N.	Ov			96.25	0.98
Glass	3/8 Rule	23	7.2	93.72	1.49
	Open N-C 2-point	23	7.3		
Ionosphere	Open N-C 7-point	8	0.3	91.29	1.49
Iris	CT, CF, CTMin, CMinT	1, 7, 8, 11, 14, 25	*	98.00	2.21
	CTT, CTC, CDiff1, Open N-C 2-point	1, 8, 11, 14, 25			
	CDiff2	1, 7, 8, 9, 11, 13, 14, 25			
Isolet	CT, CF	11, 14	*	93.34	0.66
	CTMin, CMinT	1, 11, 14			
	CDiff2	1, 8, 11, 14			
Letter R.	CT, CF, CTMin, CMinT	11, 14	*	93.49	0.91
	CDiff2	1, 11, 14			
Onehr	Extr. N-C 4-point	8	0.3	92.03	0.89
Online N.	Ov	0.6	92.01	0.62	
Page B.	Extr. N-C 3-point	14	-9.8	94.18	0.62
Predict S.	CDiff2	9	0.1	96.31	1.00
Satimage	CT, CF, CTMin, CMinT, CDiff2	1, 8, 11, 14, 25	*	93.74	0.63
	CDiff1, Open N-C 2-point	1, 11, 14			
Segment	CTC	11, 14	6.4, 6.5	93.40	0.63
Shuttle	N-C 10-point	16	4	94.50	0.83
	N-C 10-point	17	4.6, 4.7		
Skin N.	Extr. N-C 4-point	7	5.3	91.72	1.35
Spambase	Ov			94.53	0.51
Wine	CT, CF, CTMin, CMinT	1, 7, 8, 11, 14, 25	*	98.33	1.84
	CTC, CDiff1, Simpson, 3/8 Rule, Open N-C 2-point	1, 8, 11, 14, 25			
	CDiff2	1, 3, 7, 8, 9, 11, 13, 14, 25			
Wine Q.R.	CT, CF, CTMin, CMinT, CDiff2	1, 8, 11, 14, 25	*	93.63	0.60
	CTC, CDiff1	1, 11, 14			
	Open N-C 2-point	1, 8, 11, 14			
Wine Q.W.	CT, CF, CTMin, CMinT, CDiff1, CDiff2	1, 8, 11, 14, 25	*	93.41	0.83
	CTC	1, 11, 14			
	Open N-C 2-point	1, 8, 11, 14			
Zoo	CT, CF, CTMin, CMinT	1, 8, 11, 14	*	94.50	1.50
	CDiff1	11, 14			
	CDiff2, Open N-C 2-point	1, 8, 11, 14, 25			
Lung	**				

Therefore, we have used the followng datasets in the two-class problem: Mammography, Annthyroid, Pendigits, Satellite, Satimage-2, Optidigits, Thyroid, Musk, Cardio, Letter Recognition, Vowels, Pima, Arrhythmia, Ionosphere, Wbc, Vertebral, Glass, Lympho, and Wine (sorted descending by number of records) [57], [58], [59], [60], [61], [62], [63], [64], [65], [66], [67]. Their detailed properties are presented in paper [45]. We do not present them here due to lack of space.

Additionally, we have utilized Abalone, Breast Cancer Wisconsin, Electrical Grid Stability Simulated Data, Gas Sensor Array Drift Dataset at Different Concentrations, Dry Bean, EEG Eye State, Ozone Level Detection (eight and one hour peak), German Credit Data, Glass Identification, Ionosphere, Iris, Isolet, Letter Recognition, Online News Popularity, Page Blocks Classification, Students Dropout and Academic Success, Landsat Satellite, Image Segmentation, Shuttle, Skin Segmentation, Spambase, Wine, Wine Quality, Zoo, and Lung Cancer [69], [70], [71], [72], [73], [74], [75] in the multi-class classification problem. The details of the datasets used for verifying multi-class problems are shown in Table 1.

B. INDIVIDUAL CLASSIFICATION METHODS

Here, we recall the classifiers used in the experiments. When analyzing the anomaly detection-related behavior of aggregation formulae two experimental sets have been

TABLE 5. Average results of the best accuracy values obtained within the particular aggregation operators.

Choquet int. generalization	Anomaly det. acc. [%]	Classification acc. [%]
3/8 Rule	96.28	92.59
Bool	96.15	92.26
CDiff1	96.51	93.26
CDiff2	96.41	93.41
CDiff3	82.76	48.93
CF	96.57	93.49
CT	96.31	92.50
CTC	96.39	92.77
CTT	96.63	93.49
CMin	94.70	91.16
CMinM	96.57	93.22
CMMin	96.57	93.22
Durand	96.24	92.30
Extr. N-C 2-Point	96.45	92.49
Extr. N-C 3-Point	96.17	92.29
Extr. N-C 4-Point	95.72	92.03
Hardy	96.16	92.26
N-C 6-Point	96.16	92.36
N-C 7-Point	96.16	92.25
N-C 8-Point	96.13	92.25
N-C 9-Point	95.97	91.96
N-C 10-Point	96.12	92.25
N-C 11-Point	96.10	92.33
Newton-Cotes 3-point	96.05	92.20
Newton-Cotes 4-point	96.12	92.34
Open N-C 2-point	96.48	92.71
Open N-C 3-point	96.05	92.20
Open N-C 4-point	96.12	92.34
Open N-C 5-point	95.98	92.50
Open N-C 6-point	96.07	92.33
Open N-C 7-point	95.92	92.39
OV	96.20	93.32
Shovelton	96.17	92.03
Simpson	96.37	92.58
Weddle	96.16	92.24
Woolhouse	96.16	92.02

TABLE 6. Average accuracies [%] for all the CI generalizations and classical t-norms for two-class problems.

CI gen.	Prod.	Min.	Łuk.	DP	NM	Ham.
3/8 Rule	95.04	95.03	95.03	47.67	94.75	95.05
Bool	95.01	94.94	94.99	47.69	94.88	95.00
CDiff1	95.00	95.11	95.24	46.59	94.93	**95.07**
CDiff2	95.03	94.75	95.32	46.59	**95.22**	94.91
CDiff3	17.56	8.59	31.53	7.08	22.90	11.80
CF	95.02	**95.36**	**95.36**	46.59	94.85	95.19
CMinMT	94.91	94.70	95.34	46.59	94.82	94.71
CT	95.02	94.70	94.77	**48.41**	94.79	94.84
CTC	95.02	**95.36**	**95.36**	46.59	94.85	95.19
CTT	95.02	95.15	94.96	**48.41**	94.88	95.06
CTMin	94.91	94.70	95.34	46.59	94.82	94.71
Durand	95.02	94.96	94.98	47.69	94.88	95.01
Extr. N-C 2-point	**95.07**	95.01	95.06	46.59	94.86	95.05
Extr. N-C 3-point	94.80	94.59	94.99	47.36	94.86	94.67
Extr. N-C 4-point	94.67	94.57	94.66	33.82	94.63	94.58
Hardy	95.02	94.95	94.96	47.67	94.94	95.01
N-C 10-point	94.99	94.97	95.00	47.63	94.68	94.99
N-C 11-point	95.02	94.93	94.95	47.53	94.82	94.99
N-C 6-point	95.02	95.02	95.02	47.69	94.71	95.02
N-C 7-point	95.01	94.94	94.96	47.67	94.95	95.00
N-C 8-point	95.02	94.99	95.04	47.66	94.65	95.02
N-C 9-point	94.97	94.90	94.94	47.59	94.83	94.96
N-C 3-point	94.96	94.90	94.98	47.58	94.86	94.97
N-C 4-point	95.00	94.99	94.99	47.69	94.79	95.00
Open N-C 2-point	95.05	95.05	95.05	46.59	94.74	95.06
Open N-C 3-point	94.96	94.90	94.98	47.58	94.86	94.97
Open N-C 4-point	95.00	94.99	94.99	47.69	94.79	95.00
Open N-C 5-point	**95.07**	94.99	95.05	43.91	94.83	95.06
Open N-C 6-point	95.04	95.03	95.08	47.52	94.81	95.05
Open N-C 7-point	94.68	94.59	94.96	26.90	94.79	94.20
Shovelton	94.98	94.91	94.93	47.64	94.88	94.96
Simpson	95.04	95.02	95.02	47.67	94.86	95.03
Weddle	95.01	94.94	94.95	47.67	94.95	95.00
Woolhouse	94.97	94.91	94.93	47.65	94.82	94.96

realized. During the first run the default settings of PyCaret library [76] were applied for the following methods: Clustering, COF, IF, Histogram, k-NN, LOF, SVM, and PCA.

In the case of multi-class problem Naive Bayes, K-NN, Extreme Gradient Boosting (GB), Light GB, Random Forest, Extra Trees, CatBoost, GB, Ada Boost, Decision Trees, Ridge, and SVM were used. In this manner, we have conducted 19 experiments with the classifiers applied to two-class anomaly detection problem and 27 experiments with 12 classifiers in the problems with many classes. Each experiment was repeated 10 times to obtain stable results.

Here, it is worth to emphasis that any classifier returns an element membership probability to a specific class or anomaly score in the case of anomaly detection. All the resulted values can be normalized to the range [0, 1]. In this way, the probabilities of belonging to a group of outliers or anomalies as well as a group of *typical* or *normal* points can be determined. Therefore, all the aggregation models give the resulting values from [0, 1]. As the densities (weights) of the models we have proposed the average $F1$ score measures found on a basis of a few randomly chosen ODDS datasets. Similarly, the weights assigned to individual classifiers in the aggregation processes of multi-class problem were the normalized values of accuracies determined in the initial series of pretests. To be more precise, we have performed tests on a few randomly selected databases using each of the classifiers. The resulting accuracy measures serve as weights. The sum of the weights gives the value 1.

C. AGGREGATION OF INDIVIDUAL CLASSIFICATION RESULTS

Here, we thoroughly discuss the aggregation of classifiers-based results. We use the general model described by (18) as well as other formuale described in the previous sections and in the Appendix. To be more precise, we replace the t-norm $T(\cdot,\cdot)$ appearing in (18) by 25 various triangular norm groups detailed in the book [1]. All the t-norms are used with different parameters belonging to the interval $[-10, 10]$. The step is 0.1. In the cases where, pursuant to the definition of the t-norm, the range of the parameter is limited, we have narrowed down the range of the loop accordingly. Typical situations of this kind were if the parameter was positive or non-zero.

An important aspect of the results presented here should be emphasized. Namely, in many papers on the Choquet integral and its generalizations its superiority over other classical methods of aggregation was proved. Hence, we omit such considerations in this study, referring only to classes of aggregators closely related to the Choquet integral. However, this is not a narrowing down of the problem, but an attempt to focus it on this broad class of functions.

In the case of anomaly detection, the best average accuracy measure result calculated over all the considered databases, namely 95.53%, was obtained with 11-point Newton-Cotes formula being placed under the integral sign and t-norm

$$T_7(t,s) = \max(\alpha(t+s-1) + (1-\alpha)ts, 0) \tag{27}$$

TABLE 7. Average accuracies [%] for all the CI generalizations and classical t-norms for multi-class problems.

CI gen.	Prod.	Min.	Łuk.	DP	NM	Ham.
3 8 Rule	91.56	**91.74**	91.43	71.69	91.81	91.61
Bool	91.51	91.64	91.41	71.74	91.75	91.55
CDiff1	91.46	90.95	91.5	71.58	91.04	91.63
CDiff2	**91.64**	90.59	**91.83**	71.59	91.54	91.2
CDiff3	5.56	5.3	12.96	29.87	12.14	5.35
CF	91.47	90.74	91.54	71.59	91.05	**91.72**
CMinM	91.47	91.16	91.71	71.59	91.48	91.29
CMMin	91.47	91.16	91.71	71.59	91.48	91.29
CT	91.47	91.16	91.18	67.39	86.47	91.41
CTC	91.47	90.74	91.54	71.59	91.05	**91.72**
CTT	91.47	91.34	91.65	71.74	89.98	91.53
Durand	91.51	91.64	91.41	71.75	91.74	91.56
Extr. N-C 2-point	91.55	91.8	91.45	71.68	**91.89**	91.61
Extr. N-C 3-point	91.33	91.52	91.28	71.34	91.59	91.38
Extr. N-C 4-point	91.27	91.44	91.23	15.94	91.46	91.33
Hardy	91.5	91.6	91.36	71.7	91.71	91.54
N-C 10-point	91.48	91.5	91.36	71.59	91.59	91.5
N-C 11-point	91.47	91.5	91.38	71.7	91.53	91.51
N-C 6-point	91.54	91.67	91.41	71.75	91.73	91.56
N-C 7-point	91.5	91.6	91.37	71.69	91.7	91.54
N-C 8-point	91.5	91.59	91.38	71.7	91.67	91.54
N-C 9-point	91.43	91.52	91.35	71.55	91.6	91.49
N-C 3-point	91.5	91.62	91.39	71.68	91.71	91.55
N-C 4-point	91.51	91.64	91.4	**71.76**	91.71	91.56
Open N-C 2-point	91.58	91.8	91.44	71.59	91.86	91.62
Open N-C 3-point	91.5	91.62	91.39	71.68	91.71	91.55
Open N-C 4-point	91.51	91.64	91.4	71.76	91.71	91.56
Open N-C 5-point	91.53	91.67	91.41	71.66	91.76	91.59
Open N-C 6-point	91.51	91.64	91.4	71.68	91.68	91.54
Open N-C 7-point	91.3	91.38	91.31	6.96	91.7	91.39
Shovelton	91.44	91.47	91.34	71.6	91.52	91.47
Simpson	91.56	91.72	91.43	71.67	91.83	91.61
Weddle	91.5	91.59	91.37	71.7	91.71	91.54
Woolhouse	91.44	91.47	91.34	71.59	91.53	91.47

with $\alpha = 3.5$, see Table 2. Also other quadrature-inspired Choquet integral generalizations are good choices here, in particular

$$T_{11}(t,s) = \left(\max\left(t^\alpha s^\alpha - 2\left(1-t^\alpha\right)\left(1-s^\alpha\right),0\right)\right)^{\frac{1}{\alpha}}, \alpha > 0 \tag{28}$$

$$T_{14}(t,s) = \left(\max\left(\frac{t^\alpha s^\alpha - 2\left(1-t^\alpha\right)\left(1-s^\alpha\right)}{1-\left(1-t^\alpha\right)\left(1-s^\alpha\right)},0\right)\right)^{\frac{1}{\alpha}} \tag{29}$$

where $\alpha \neq 0$, and

$$T_{25}(t,s) = \begin{cases} \left(1-\left(1-t^\alpha\right)\sqrt{1-\left(1-s^\alpha\right)^2} \right. \\ \left. -\left(1-s^\alpha\right)\sqrt{1-\left(1-t^\alpha\right)^2}\right)^{\frac{1}{\alpha}} & , \alpha > 0 \\ \text{if} \quad \left(1-t^\alpha\right)^1 + \left(1-s^\alpha\right)^1 \leq 1 \\ 0 \quad \text{otherwise} \end{cases} \tag{30}$$

Fig. 1 depicts the ROC curves with AUC values for the best generalizations of the Choquet integral which are specified in Table 2 for anomaly detection task. The graphs were obtained using the Annthyroid database, for the parameters as in the table. We supplemented them with plots of the C_{Diff_2} function, which yielded the best average measure of accuracy for this dataset and the classical Choquet integral. Indeed, C_{Diff_1} and C_{Diff_2} have the highest AUC measures. Moreover, a good result was obtained for the Newton-Cotes 11-point quadrature. This shows that it is worth using proposed kinds of generalizations, especially because the classical Choquet integral performs average in this comparison.

In Tables 3, 4, and 5 listed are the best results obtained for each dataset. Considering the two-class anomaly detection problem it is interesting to observe that the t-norm no. 23 is dominating, i.e. it wins in the most cases. It reads as follows

$$T_{23}(t,s) = \left(\log\left(e^{t^{-\alpha}} + e^{s^{-\alpha}} - e\right)\right)^{-\frac{1}{\alpha}}, \alpha > 0 \tag{31}$$

There is also one more interesting observation here. Namely, if we divide the Choquet integral generalizations onto two groups, namely quadrature-inspired extensions and the ones presented in the literature of the topic, the representatives of the first class of functions give better results here in terms of number of datasets, where such kind of integral result in the best accuracy. There are many various winners in this group, for instance, Newton-Cotes (N-C), extrapolative N-C, or open N-C formulae. Hence, each dataset is worth examining separately to design the best parameters. However, the above T_{23} norm is worth considering at the beginning in each case. Note that the asterisk symbol * means various parameters. We do not report the details since many t-norms and their parameters that differ dependent on the case. Moreover, in the case of Lung Cancer dataset (denoted by **), too many various classifiers have won and there is no sense to point them out here.

Now, let us consider the multi-class problems, i.e. the problems solved by typical machine learning classifiers, see Table 4. Here, there are a few problems for which many various functions return the same top average result. Also classic extensions of Choquet integral give a few top results more than the extensions inspired by quadratures. However, the latter are worth considering in a more comprehensive way. Beside the t-norms no. 7, 11, 14, and 25, two another t-norms are worth merging with the Choquet-based integrals. They are

$$T_1(t,s) = \left(\max\left(t^{-\alpha} + s^{-\alpha}\right),0\right)^{-\frac{1}{\alpha}}, \alpha \neq 0 \tag{32}$$

and

$$T_8(t,s) = \frac{\left(\alpha^2 ts - (1-t)(1-s),0\right)}{\alpha^2 - (\alpha-1)^2(1-t)(1-s)}, \alpha \geq 0 \tag{33}$$

Typical generalization of the CI with a t-norm replacing the product and simple modifications of it give good results here. However, very promising results are returned by open Newton-Cotes 2-point formula-based Choquet integral as well as extrapolative N-C 4-point formula-based CI. There is one dataset (Lung Cancer) for which almost all the aggregation functions returned 100% recognition rate. Therefore, the results are omitted in the table (it is denoted **).

Fig. 2 depicts the average accuracies obtained for various classifiers with the functions T_1, T_8, T_{11}, T_{14}, and T_{25} substituting the T in the case of open Newton-Cotes 2-point quadrature given by (36), see Appendix. The first three of the t-norms are performing well for a positive range of α parameter values while the last two of them return good results for α values being close to 0. However, the plot confirms the assumption that it is relatively difficult to choose the right value of the α parameter, but it can be estimated on the basis of experience or general statistics presented in this study. On the other hand, if we consider T_1 and T_8 t-norms we see that the parameter λ with relatively large values is an optimal choice. Other t-norms depicted in Fig. 2 perform well when λ tends to zero point. It may be related with the formulae (28), (29), (30), (32), and (33).

Now, let us discuss the average values list of the best results produced within each type of aggregation functions over all the cases analyzed in the tests with respect to anomaly detection and classification task. Table 5 presents the standing. Note that here any divisions into t-norms and their parameters are not considered. We take into account only the best result obtained by each extension of Choquet integral for specific datasets. Therefore, it is hard to say how much it is helpful. However, it shows that some of the CI extensions are quite stable in the sense of accuracies. In the case of anomaly detection problem all the functions give similar results except the worse results obtained by C_{Diff_3}, C_{Min}, 9-point N-C, 7 point open N-C, or 9-point open N-C quadrature-inspired formulae. On the other hand, in the case of multi-class problem a few generalizations are leading, namely C_{Diff_1}, C_{Diff_2}, C_T, C_F, C_{MinT}, C_{TMin}, or pre-aggregation function with overlapping formula under the integral sign. This overlapping function is

$$t(t,s)=\begin{cases}\dfrac{ts}{(p+(1-p)(t+s-ts))} & \text{for } \dfrac{ts}{(p+(1-p)(t+s-ts))}<\alpha(*)\\ \alpha & \text{if not } (*) \text{ and } \dfrac{ts}{(p+(1-p)(t+s-ts))}<\beta\\ \alpha+\dfrac{ts}{(p+(1-p)(t+s-ts))(1-\beta)} & \text{otherwise}\end{cases} \quad (34)$$

Now, we analyze the CI extensions performance when only the typical kinds of t-norms are considered, see e.g. [25]. They are minimum, algebraic product (note that C_F and C_T are simple Choquet integral here), Łukasiewicz t-norm, drastic product (DP), nilpotent minimum (NM), and Hamacher t-norm given by (21)-(26), respectively. Tables 6 and 7 reveal an interesting fact. It is quite inverse situation than in the previous case. Namely, in the two-class problem when typical t-norms cooperate well with classic extensions of the Choquet integral (C_F, C_T, C_{TC}, C_{TT}) while in the multi-class task the numerical methods-inspired CI generalizations are slightly better (C_{Diff2}, extrapolative N-C 2-point, or 3/8 rule).

Finally, one can observe that the typical CI gives a relatively good results, but not the best, i.e. average accuracies at the level of 95.02% and 91.47% for multi-class and two-class problems, respectively.

Several observations can be drawn from the above series of experiments and their results. Namely, it is impossible to indicate one universal aggregation method (in the sense of the final set of its parameters) for all possible cases. However, it is worth stressing that in the case of big data sets, and when a relatively large classifier count is used, searching for such a method based on tests for all possible cases (e.g. 25 families of t-norms, ranges of several hundred values of their parameters, various variants of quadrature and different forms CI generalizations) can cause a waste of time and hardware resources that practitioners cannot afford. Therefore, this study is a kind of guide and an indication among which sets of parameters are certainly worth looking for, and which can certainly be omitted. Such parameter combinations have been presented in this experimental section.

Moreover, these experiments prove that it is very easy to make a mistake. It is very difficult to raise the best result, but it is easy to choose parameters for which the quality of classification/anomaly detection drops drastically.

VI. CONCLUSION AND FUTURE WORK

In this study, we have thoroughly discussed a novel approach to determine the Choquet integral's application in crucial decision-making problems, specifically in aggregating classifier results. This approach involves substituting the triangular norm, typically found in the Choquet integral, with a formula based on a foundation of numerical quadratures. We've combined 25 families of triangular norms, each encompassing a wide range of parameters, with over 30 enhancements of CI, some derived from existing literature and others being new proposals. These newly introduced approaches have demonstrated their efficiency and value across 27 multi-class and 19 two-class problems.

We have introduced an extensive analysis of diverse integrand variations found in CI generalizations, focusing on their efficiency as aggregation functions. We have presented parameter configurations that maximize the probability of achieving the best classification results. These optimal configurations apply to various databases and problem classes. Additionally, we've identified which Choquet integral generalizations, based on quadrature methods, are effective and compatible with specific t-norms. The results obtained through a series of numerical experiments have demonstrated that our approach can outperform previous CI-based aggregation methods or, at the very least, achieve a similar level of efficiency in terms of accuracy.

In the future, it is worth to build another CI generalizations using the differences between the quadratures appearing after the integral symbol, namely

$$Ch_{QQ}=\sum_{i=1}^{n}\left(T\left(Q\left(h_i\right)-Q\left(h_{i+1}\right),m\left(A_i\right)\right)\right) \quad (35)$$

Also, it is worth to apply the models proposed in this study to specific domains of applications such as datasets coming from medical diagnosis, questionnaires, industrial measurements, etc. Moreover, not only the accuracy measures can be aggregated. For instance, one can examine such kind of operators in a process of aggregation of larger number of classifiers, for instance the results obtained by each of 100 trees in the Isolation Forest anomaly detection technique [58]. Despite the general properties are quite known, a deeper examining of a choice of various t-norms and their parameters is still valuable task. Finally, it is worth stressing that similar techniques can be applied to Ordered Weighted Averaging Operators (OWA) and their generalizations.

APPENDIX
QUADRATURE FORMULAE

Here, we recall chosen expressions inspired by quadrature rules. For the symbol notions, see Section III. The formulae are (see [53]) trapezoidal (open Newton-Cotes 2-point) rule

$$Q_T(h_i) = \frac{1}{2}(h_i + h_{i+1}) \tag{36}$$

3/8 rule (called Simpson formula)

$$Q_{3/8}(h_i) = \frac{1}{8}(h_{i-1} + 3h_i + 3h_{i+1} + h_{i+2}) \tag{37}$$

3/8 rule (Bool's)

$$Q_B(h_i) = \frac{1}{90}(7h_{i-2} + 32h_{i-1} + 12h_i + 32h_{i+1} + 7h_{i+2}) \tag{38}$$

Hardy's rule

$$Q_H(h_i) = \frac{1}{600}(28h_{i-3} + 162h_{i-2} + 12h_i + 162h_{i+2} + 28h_{i+3}) \tag{39}$$

Durand's rule

$$Q_D(h_i) = \frac{1}{40}(4h_{i-2} + 11h_{i-1} + 10h_i + 11h_{i+1} + 4h_{i+2}) \tag{40}$$

a version of Shovelton's rule

$$Q_{Sh}(h_i) = \frac{1}{252}(8h_{i-5} + 35h_{i-4} + 15h_{i-3} + 35h_{i-2} + 15h_{i-1} + 36h_i + 15h_{i+1} + 35h_{i+2} + 15h_{i+3} + 35h_{i+4} + 8h_{i+5}) \tag{41}$$

Weddle's rule form

$$Q_{We}(h_i) = \frac{1}{20}(h_{i-3} + 5h_{i-2} + h_{i-1} + 6h_i + h_{i+1} + 5h_{i+2} + h_{i+3}) \tag{42}$$

Woolhouse's rule

$$Q_{Wo}(h_i) = \frac{1}{2}\left(\frac{223}{3969}h_{i-5} + \frac{5875}{18144}h_{i-4} + \frac{4625}{10584}h_{i-2} + \frac{41}{112}h_i + \frac{4625}{10584}h_{i+2} + \frac{5875}{18144}h_{i+4} + \frac{223}{3969}h_{i+5}\right) \tag{43}$$

Newton-Cotes (N-C) 6-point rule

$$Q_{NC6}(h_i) = \frac{1}{288}(19h_{i-2} + 75h_{i-1} + 50h_i + 50h_{i+1} + 75h_{i+2} + 19h_{i+3}) \tag{44}$$

N-C 7-point rule

$$Q_{NC7}(h_i) = \frac{1}{840}(41h_{i-3} + 216h_{i-2} + 27h_{i-1} + 272h_i + 27h_{i+1} + 216h_{i+2} + 41h_{i+3}) \tag{45}$$

N-C 8-point rule

$$Q_{NC8}(h_i) = \frac{1}{17280}(751h_{i-3} + 3577h_{i-2} + 1323h_{i-1} + 2989h_i + 2989h_{i+1} + 1323h_{i+2} + 3577h_{i+3} + 751h_{i+4}) \tag{46}$$

N-C 9-point rule

$$Q_{NC9}(h_i) = \frac{1}{28350}(989h_{i-4} + 5888h_{i-3} - 928h_{i-2} + 10496h_{i-1} - 4540h_i + 10496h_{i+1} - 928h_{i+2} + 5888h_{i+3} + 989h_{i+4}) \tag{47}$$

N-C 10-point rule

$$Q_{NC10}(h_i) = \frac{1}{89060}(2857h_{i-4} + 15741h_{i-3} + 1080h_{i-2} + 19344h_{i-1} + 5778h_i + 5778h_{i+1} + 19344h_{i+2} + 1080h_{i+3} + 15741h_{i+4} + 2857h_{i+5}) \tag{48}$$

Newton-Cotes 11-point rule.

$$Q_{NC11}(h_i) = \frac{1}{598752}(16067h_{i-5} + 106300h_{i-4} - 48525h_{i-3} + 272400h_{i-2} - 260550h_{i-1} + 427368h_i - 260550h_{i+1} + 272400h_{i+2} - 48525h_{i+3} + 106300h_{i+4} + 16067h_{i+5}) \tag{49}$$

These formulae are called closed rules. There are also so-called open formulae. They are Newton-Cotes 3-point rule

$$Q_{ONC3}(h_i) = \frac{1}{3}(2h_{i-1} - h_i + 2h_{i+1}) \tag{50}$$

open N-C 4-point rule

$$Q_{ONC4}(h_i) = \frac{1}{24}(11h_{i-1} + h_i + h_{i+1} + 11h_{i+2}) \tag{51}$$

open N-C 5-point rule

$$Q_{ONC5}(h_i) = \frac{1}{24}(11h_{i-2} - 14h_{i-1} + 26h_i - 14h_{i+1} + 11h_{i+2}) \tag{52}$$

open N-C 6-point rule

$$Q_{ONC6}(h_i) = \frac{1}{1440}(611h_{i-2} - 453h_{i-1} + 562h_i + 562h_{i+1} - 453h_{i+2} + 611h_{i+3}) \quad (53)$$

open N-C 7-point rule

$$Q_{ONC7}(h_i) = \frac{1}{945}(460h_{i-3} - 954h_{i-2} + 2196h_{i-1} - 2459h_i + 2196h_{i+1} - 954h_{i+2} + 460h_{i+3}) \quad (54)$$

The last group of formulae are the extrapolative Newton-Cotes quadratures. In numerical analysis theory they are used to calculate an integral in an interval using the points laying around it. They are extrapolative N-C 2-point rule

$$Q_{ENC2}(h_i) = \frac{1}{3}(3h_i - h_{i+1}) \quad (55)$$

extrapolative N-C 3-point rule

$$Q_{ENC3}(h_i) = \frac{1}{12}(23h_{i-1} - 16h_i + 5h_{i+1}) \quad (56)$$

and, finally, extrapolative N-C 4-point rule

$$Q_{ENC4}(h_i) = \frac{1}{32}(55h_{i-1} - 59h_i + 37h_{i+1} - 9h_{i+2}) \quad (57)$$

REFERENCES

[1] C. Alsina, M. J. Frank, and B. Schweizer, *Associative Functions. Triangular Norms and Copulas*. Singapore: World Scientific, 2006.
[2] M. Baczyński, H. Bustince, and R. Mesiar, "Aggregation functions: Theory and applications," *Fuzzy Set. Syst.*, vol. 324, p. 325, May 2017.
[3] G. Beliakov, A. Pradera, and T. Calvo, *Aggregation Functions: A Guide for Practitioners*. Berlin, Germany: Springer–Verlag, 2007.
[4] T. Calvo, G. Mayor, and R. Mesiar, *Aggregation Operators. New Trends and Applications*. Heidelberg, Germany: Physica-Verlag, 2014.
[5] M. Gagolewski, *Data Fusion: Theory, Methods, and Applications* (Institute of Computer Science). Warsaw, Poland: Polish Academy of Sciences, 2015.
[6] M. Grabisch, J. Marichal, R. Mesiar, and E. Pap, *Aggregation Functions (Encyclopedia of Mathematics and Its Applications)*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[7] R. Mesiar, A. Kolesárová, T. Calvo, and M. Komorníková, "A review of aggregation functions," in *Fuzzy Sets and Their Extensions: Representation, Aggregation and Models*, vol. 220, H. Bustince, F. Herrera, and J. Montero, Eds. Springer, 2008, pp. 121–144.
[8] M. Grabisch, J.-L. Marichal, R. Mesiar, and E. Pap, "Aggregation functions: Means," *Inf. Sci.*, vol. 181, no. 1, pp. 1–22, Jan. 2011.
[9] M. Grabisch, J.-L. Marichal, R. Mesiar, and E. Pap, "Aggregation functions: Construction methods, conjunctive, disjunctive and mixed classes," *Inf. Sci.*, vol. 181, no. 1, pp. 23–43, Jan. 2011.
[10] R. Pérez-Fernández and B. De Baets, "Aggregation theory revisited," *IEEE Trans. Fuzzy Syst.*, vol. 29, no. 4, pp. 797–804, Apr. 2021.
[11] F. Blanco-Mesa, E. León-Castro, and J. M. Merigó, "A bibliometric analysis of aggregation operators," *Appl. Soft Comput.*, vol. 81, Aug. 2019, Art. no. 105488.
[12] E. P. Klement, R. Mesiar, and E. Pap, *Triangular Norms*. Dordrecht, The Netherlands: Springer, 2000.
[13] E. P. Klement and R. Mesiar, *Logica, Algebraic, Analytic, and Probabilistic Aspects of Triangular Norms*. Amsterdam, The Netherlands: Elsevier, 2005.
[14] R. R. Yager and J. Kacprzyk, *The Ordered Weighted Averaging Operators: Theory and Applications*. New York, NY, USA: Springer, 2012.
[15] L. Jin, R. Mesiar, and R. Yager, "Ordered weighted averaging aggregation on convex poset," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 3, pp. 612–617, Mar. 2019.
[16] G. Beliakov, T. Cao, and V. Mak-Hau, "Aggregation of interacting criteria in land combat vehicle selection by using fuzzy measures," *IEEE Trans. Fuzzy Syst.*, vol. 30, no. 9, pp. 3979–3989, Sep. 2022.
[17] P. R. Hemantkumar, T. Gupta, A. Vashisth, P. Chintha, A. V. Wargiya, M. G. George, M. A. Alexander, and S. J. Narayanan, "GCN embedded with polynomial aggregation function for group activity recognition," in *Proc. IEEE 4th Conf. Inf. Commun. Technol. (CICT)*, Dec. 2020, pp. 1–6.
[18] W. Pedrycz, A. Gacek, and X. Wang, "Aggregation of order-2 fuzzy sets," *IEEE Trans. Fuzzy Syst.*, vol. 29, no. 11, pp. 3570–3575, Nov. 2021.
[19] J. M. Merigó, A. M. Gil-Lafuente, D. Yu, and C. Llopis-Albert, "Fuzzy decision making in complex frameworks with generalized aggregation operators," *Appl. Soft Comput.*, vol. 68, pp. 314–321, Jul. 2018.
[20] B. Zhang, W. Pedrycz, A. R. Fayek, A. Gacek, and Y. Dong, "Granular aggregation of fuzzy rule-based models in distributed data environment," *IEEE Trans. Fuzzy Syst.*, vol. 29, no. 5, pp. 1297–1310, May 2021.
[21] D. Rutkowska, D. Kurach, and E. Rakus-Andersson, "Fuzzy granulation approach to face recognition," in *Artificial Intelligence and Soft Computing* (Lecture Notes in Computer Science), vol. 12855, L. Rutkowski, R. Scherer, M. Korytkowski, W. Pedrycz, R. Tadeusiewicz, J. M. Zurada, Eds. Cham, Switzerland: Springer, 2021, pp. 495–510.
[22] G. Choquet, "Theory of capacities," in *Proc. Annales de l'Institut Fourier*, vol. 5, 1953, pp. 131–295.
[23] M. Grabisch, "The application of fuzzy integrals in multicriteria decision making," *Eur. J. Oper. Res.*, vol. 89, no. 3, pp. 445–456, Mar. 1996.
[24] H. Bustince, J. A. Sanz, G. Lucca, G. P. Dimuro, B. Bedregal, R. Mesiar, A. Kolesárová, and G. Ochoa, "Pre-aggregation functions: Definition, properties and construction methods," in *Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ-IEEE)*, Jul. 2016, pp. 294–300.
[25] G. Lucca, J. A. Sanz, G. P. Dimuro, B. Bedregal, R. Mesiar, A. Kolesárová, and H. Bustince, "The notion of pre-aggregation function," in *Proc. MDAI LNAI*, V. Torra and Y. Narakawa, Eds. 2015, pp. 33–41.
[26] H. Bustince, R. Mesiar, J. Fernandez, M. Galar, D. Paternain, A. Altalhi, G. P. Dimuro, B. Bedregal, and Z. Takác, "d-Choquet integrals: Choquet integrals based on dissimilarities," *Fuzzy Sets Syst.*, vol. 414, pp. 1–27, Jul. 2021.
[27] P. Karczmarek, A. Kiersztyn, and W. Pedrycz, "Generalized Choquet integral for face recognition," *Int. J. Fuzzy Syst.*, vol. 20, no. 3, pp. 1047–1055, Mar. 2018.
[28] G. P. Dimuro, J. Fernández, B. Bedregal, R. Mesiar, J. A. Sanz, G. Lucca, and H. Bustince, "The state-of-art of the generalizations of the Choquet integral: From aggregation and pre-aggregation to ordered directionally monotone functions," *Inf. Fusion*, vol. 57, pp. 27–43, May 2020.
[29] G. Beliakov, S. James, A. Kolesárová, and R. Mesiar, "Cardinality-limiting extended pre-aggregation functions," *Inf. Fusion*, vol. 76, pp. 66–74, Dec. 2021.
[30] P. Karczmarek, L. Galka, A. Kiersztyn, M. Dolecki, K. Kiersztyn, and W. Pedrycz, "Choquet integral-based aggregation for the analysis of anomalies occurrence in sustainable transportation systems," *IEEE Trans. Fuzzy Syst.*, vol. 31, no. 2, pp. 536–546, Feb. 2023.
[31] P. Drygas, B. Pekala, K. Balicki, and D. Kosior, "Influence of new interval-valued pre-aggregation function on medical decision making," in *Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ-IEEE)*, Jul. 2020, pp. 1–8.
[32] E. Hinojosa-Cardenas, E. Sarmiento-Calisaya, H. A. Camargo, and J. A. Sanz, "Improving Michigan-style fuzzy-rule base classification generation using a Choquet-like Copula-based aggregation function," in *Proc. WILF 13th Int. Workshop Fuzzy Log. Appl.*, 2021, pp. 1–9.
[33] T. Batista, B. Bedregal, and R. Moraes, "Constructing multi-layer classifier ensembles using the choquet integral based on overlap and quasi-overlap functions," *Neurocomputing*, vol. 500, pp. 413–421, Aug. 2022.
[34] M. R. Hashmi and M. Riaz, "Diagnosis of lumbar degenerative disc disease by using L^p-spaces related to generalized interval-valued *m*-polar neutrosophic Choquet integral operator," *Int. J. Biomathematics*, vol. 14, no. 8, Nov. 2021, Art. no. 2150063.
[35] L. Sha and Y. Shao, "Fermatean hesitant fuzzy Choquet integral aggregation operators," *IEEE Access*, vol. 11, pp. 38548–38562, 2023.
[36] P. Karczmarek, W. Pedrycz, A. Kiersztyn, and M. Dolecki, "A comprehensive experimental comparison of the aggregation techniques for face recognition," *Iran. J. Fuzzy Syst.*, vol. 16, no. 4, pp. 1–19, 2019.

[37] S.-P. Wan, J. Yan, W.-C. Zou, and J.-Y. Dong, "Generalized Shapley Choquet integral operator based method for interactive interval-valued hesitant fuzzy uncertain linguistic multi-criteria group decision making," *IEEE Access*, vol. 8, pp. 202194–202215, 2020.

[38] H. Song, Z. Gong, G. Wei, W. Guo, X. Ma, and E. Herrera-Viedma, "Non-additive robust ordinal regression: A multicriteria decision model based on hierarchical-level-bidirectional choquet integral," *Appl. Soft Comput.*, vol. 146, Oct. 2023, Art. no. 110643.

[39] S. Rahnamay Bonab, S. Jafarzadeh Ghoushchi, M. Deveci, and G. Haseli, "Logistic autonomous vehicles assessment using decision support model under spherical fuzzy set integrated Choquet integral approach," *Exp. Syst. Appl.*, vol. 214, Mar. 2023, Art. no. 119205.

[40] P. Karczmarek, *Selected Problems of Face Recognition and Decision-Making Theory*. Lublin, Poland: Lublin University of Technology Press, 2018.

[41] T. S. Bisht, D. Kumar, and B. J. Alappat, "Selection of optimal aggregation function for the revised leachate pollution index (r-LPI)," *Environ. Monitor. Assessment*, vol. 194, no. 3, p. 187, Mar. 2022.

[42] C. Labreuche, "A formal justification of a simple aggregation function based on criteria and rank weights," in *Proc. From Multiple Criteria Decis. Aid Preference Learn. (DA2PL)*, 2018, pp. 1–8.

[43] J. Lin, Q. Zhang, and F. Meng, "An approach for facility location selection based on optimal aggregation operator," *Knowl.-Based Syst.*, vol. 85, pp. 143–158, Sep. 2015.

[44] J. Lin, Q. Zhang, and F. Meng, "A novel algorithm for group decision making based on continuous optimal aggregation operator and Shapley value," *Int. J. Uncertainty, Fuzziness Knowl.-Based Syst.*, vol. 27, no. 6, pp. 969–1002, Dec. 2019.

[45] P. Karczmarek, M. Dolecki, P. Powroznik, L. Galka, W. Pedrycz, and D. Czerwinski, "Quadrature-inspired generalized Choquet integral," in *Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ-IEEE)*, Jul. 2022, pp. 1–7.

[46] P. Karczmarek, W. Pedrycz, M. Reformat, and E. Akhoundi, "A study in facial regions saliency: A fuzzy measure approach," *Soft Comput.*, vol. 18, no. 2, pp. 379–391, Feb. 2014.

[47] G. Lucca, R. R. de Vargas, G. P. Dimuro, J. A. Sanz, H. Bustince, and B. R. C. Bedregal, "Analysing some T-norm-based generalizations of the Choquet integral for different fuzzy measures with an application to fuzzy rule-based classification systems," in *Proc. ENIAC Encontro Nac. Intelig. Artif. Computacional.* São Carlos: SBC, 2014, pp. 508–513.

[48] G. Lucca, J. A. Sanz, G. P. Dimuro, B. Bedregal, M. J. Asiain, M. Elkano, and H. Bustince, "CC-integrals: Choquet-like copula-based aggregation functions and its application in fuzzy rule-based classification systems," *Knowl.-Based Syst.*, vol. 119, pp. 32–43, Mar. 2017.

[49] G. Lucca, J. Antonio Sanz, G. P. Dimuro, B. Bedregal, H. Bustince, and R. Mesiar, "CF -integrals: A new family of pre-aggregation functions with application to fuzzy rule-based classification systems," *Inf. Sci.*, vol. 435, pp. 94–110, Apr. 2018.

[50] G. P. Dimuro, G. Lucca, J. A. Sanz, H. Bustince, and B. Bedregal, "CMin-integral: A Choquet-like aggregation function based on the minimum t-norm for applications to fuzzy rule-based classification systems," in *Aggregation Functions in Theory and in Practice* (Advances in Intelligent Systems and Computing), vol. 581, V. Torra Ed. Springer, 2018, pp. 83–95.

[51] G. Lucca, J. A. Sanz, G. P. Dimuro, B. Bedregal, and H. Bustince, "Pre-aggregation functions constructed by CO-integrals applied in classification problems," in *Proc. IV CBSF*, 2016, pp. 1–11.

[52] P. Karczmarek, A. Kiersztyn, and W. Pedrycz, "Generalizations of aggregation functions for face recognition," in *Proc. ICAISC LNAI*, L. Rutkowski Eds. 2019, pp. 182–192.

[53] E. W. Weisstein, *Newton-Cotes Formulas* (MathWorld-A Wolfram Web Resource). Wolfram. Accessed: Feb. 14, 2022.

[54] M. Dravinski and T. K. Mossessian, "On evaluation of the green functions for harmonic line loads in a viscoelastic half space," *Int. J. Numer. Methods Eng.*, vol. 26, no. 4, pp. 823–841, Apr. 1988.

[55] B. Shizgal, "Numerical evaluation of integrals and derivatives," in *Spectral Methods in Chemistry and Physics* (Scientific Computation). Dordrecht, The Netherlands: Springer, 2015, pp. 109–186.

[56] H. Bustince, J. Fernandez, R. Mesiar, J. Montero, and R. Orduna, "Overlap functions," *Nonlinear Anal. Theory Methods Appl.*, vol. 72, nos. 3–4, pp. 1488–1499, 2010.

[57] N. Abe, B. Zadrozny, and J. Langford, "Outlier detection by active learning," in *Proc. KDD 12th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2006, pp. 504–509.

[58] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation forest," in *Proc. 8th IEEE Int. Conf. Data Mining*, Dec. 2008, pp. 413–422.

[59] K. M. Ting, J. T. S. Chuan, and F. T. Liu, "MASS: A new ranking measure for anomaly detection," Gippsland School of Information Technology, Morwell, VIC, Australia, Tech. Rep. 2009/1, 2009.

[60] F. Keller, E. Muller, and K. Bohm, "HiCS: High contrast subspaces for density-based outlier ranking," in *Proc. IEEE 28th Int. Conf. Data Eng.*, Washington, DC, USA, Apr. 2012, pp. 1037–1048.

[61] C. C. Aggarwal and S. Sathe, "Theoretical foundations and algorithms for outlier ensembles," *ACM SIGKDD Explorations Newslett.*, vol. 17, no. 1, pp. 24–47, Sep. 2015.

[62] S. Sathe and C. Aggarwal, "LODES: Local density meets spectral outlier detection," in *Proc. SIAM Int. Conf. Data Mining*, Jun. 2016, pp. 171–179.

[63] B. Micenkova, B. McWilliams, and I. Assent, "Learning outlier ensembles: The best of both worlds—Supervised and unsupervised," in *Proc. ACM SIGKDD Workshop ODD2: Outlier Detection Description Under Data Diversity*, 2014, pp. 1–4.

[64] S. Rayana, "ODDS library," Dept. Comput. Sci., Stony Brook Univ., Stony Brook, NY, Tech. Rep., 2016. [Online]. Available: http://odds.cs.stonybrook.edu

[65] S. Rayana and L. Akoglu, "Less is more: Building selective anomaly ensembles with application to event detection in temporal graphs," in *Proc. SIAM Int. Conf. Data Mining*, Jun. 2015, pp. 622–630.

[66] A. Lazarevic and V. Kumar, "Feature bagging for outlier detection," in *Proc. 11th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2005, pp. 157–166.

[67] A. Zimek, M. Gaudet, R. J. Campello, and J. Sander, "Subsampling for efficient and effective unsupervised outlier detection ensembles," in *Proc. KDD 19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2013, pp. 428–436.

[68] D. Dua and C. Graff, "UCI machine learning repository," School Inf. Comput. Sci., Univ. California, Irvine, CA, USA, Tech. Rep., 2019. [Online]. Available: http://archive.ics.uci.edu/ml

[69] A. Vergara, S. Vembu, T. Ayhan, M. Ryan, M. Homer, and R. Huerta, "Chemical gas sensor drift compensation using classifier ensembles," *Sens. Actuators B, Chem.*, vol. 166, pp. 320–329, May 2012.

[70] I. Rodriguez-Lujan, J. Fonollosa, A. Vergara, M. Homer, and R. Huerta, "On the calibration of sensor arrays for pattern recognition using the minimal number of experiments," *Chemometric Intell. Lab. Syst.*, vol. 130, pp. 123–134, Jan. 2014.

[71] M. Koklu and I. A. Ozkan, "Multiclass classification of dry beans using computer vision and machine learning techniques," *Comput. Electron. Agricult.*, vol. 174, Jul. 2020, Art. no. 105507.

[72] K. Fernandes, P. Vinagre, and P. Cortez, "Support system for predicting the popularity of online news," in *Proc. 17th EPIA Portuguese Conf. Artif. Intell.*, 2015, pp. 535–546.

[73] M. V. Martins, D. Tolledo, J. Machado, L. M. T. Baptista, and V. Realinho, "Early prediction of student's performance in higher education: A case study," in *Proc. Trends Appl. Inf. Syst. Technol.*, vol. 1, 2021, pp. 166–175.

[74] R. B. Bhatt, G. Sharma, A. Dhall, and S. Chaudhury, "Efficient skin region segmentation using low complexity fuzzy decision tree model," in *Proc. Annu. IEEE India Conf.*, 2009, pp. 1–4.

[75] P. Cortez, A. Cerdeira, F. Almeida, T. Matos, and J. Reis, "Modeling wine preferences by data mining from physicochemical properties," *Decis. Support Syst.*, vol. 47, no. 4, pp. 547–553, Nov. 2009.

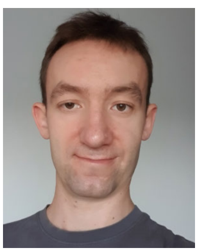
[76] A. Moez. (2020). *PyCaret: An open source, low-code machine learning library in Python*. [Online]. Available: https://pycaret.org/about.PyCaretversion1.0.0

PAWEŁ KARCZMAREK received the Ph.D. degree in mathematics from the University of Gdałsk, Poland, in 2010, and the Habilitation (D.Sc.) degree in computer science from the Systems Research Institute, Polish Academy of Sciences, in 2019. He is currently a Professor with the Lublin University of Technology, Poland, and the Head of the Department of Computational Intelligence. He is the author of more than 80 research articles and one monograph. His current research interests include computational intelligence, anomaly and deep fake detection, multi-criteria decision-making theory, and pattern recognition.

MICHAŁ DOLECKI graduated in computer science and mathematics from The John Paul II Catholic University of Lublin. He received the Ph.D. degree in computer science from the Faculty of Automatic Control, Electronics and Computer Science, Silesian University of Technology, Gliwice. He is currently an Assistant Professor with the Department of Computer Science, Lublin University of Technology. His current research interests include the various uses of biologically inspired artificial intelligence methods like neural networks and genetic algorithms. In particular, he applies them to improve the software testing process, general pattern recognition, anomaly detection, and cryptography.

PAWEŁ POWROZNIK (Member, IEEE) received the B.S. and M.S. degrees in computer science from Maria Curie-Sklodowska University (UMCS), Lublin, Poland, in 2009 and 2011, respectively, and the Ph.D. degree in electrical engineering from the Lublin University of Technology (LUT), Lublin, in 2018. He joined the Department of Computer Science, LUT, in 2016, where he is currently an Assistant Professor. Since 2011, he has been working as an iOS Mobile Application Developer. He is the author and the coauthor of scientific articles, published in journals, books, and monographs, and presented at national and international conferences. His current research interests include graph neural networks, classification, and prediction issues as well as signal processing or motion analysis topics.

ZBIGNIEW A. ŁAGODOWSKI received the degree in applied mathematics from the Faculty of Fundamental Technological Problems, AGH University of Science and Technology, Krakow, the Ph.D. degree in mathematical sciences from the University of Maria Curie-Sklodowska, Lublin, and the Habilitation degree from the University of Wroclaw. During the master's thesis, he focused on the stability of stochastic dynamical systems. After the M.S. degree, he continued his research in probability theory. In 2019, he completed a monograph on the limiting properties of random fields. He is currently a Professor with the Lublin University of Technology and the Head of the Mathematics Department. He is the author of about 40 scientific publications and the author and the editor of two monographs. His current research interests include the problems of mathematical modeling in agrophysics and computational intelligence.



ADAM GREGOSIEWICZ received the Ph.D. degree in mathematics from the University of Gdałsk, Poland. He is currently an Assistant Professor with the Department of Mathematics, Lublin University of Technology, Poland. His current research interests include functional analysis (operator theory, strongly continuous semigroups, and cosine functions), mathematical modeling in biology and physics, and recently also includes combinatorics and its connections to computer science.

ŁUKASZ GAŁKA received the M.Sc. (Eng.) degree in computer science from the Faculty of Electrical Engineering and Computer Science, Lublin University of Technology, Poland, in 2021. Currently, he is an Assistant Researcher with the Department of Computer Science, Lublin University of Technology. His current research interests include anomaly detection, data mining, and artificial intelligence in computer games.

WITOLD PEDRYCZ (Life Fellow, IEEE) is currently a Professor and the Canada Research Chair (CRC) of computational intelligence with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada. He is also with the Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland. His current research interests include computational intelligence, fuzzy modeling and granular computing, knowledge discovery and data science, pattern recognition, data science, knowledge-based neural networks, and control engineering. He has published several in these areas; the current H-index is 109 (Google Scholar) and 82 on the list top-h scientists for computer science and electronics http://www.guide2research.com/scientists/. He is also the author of 18 research monographs and edited volumes covering various aspects of computational intelligence, data mining, and software engineering.

In 2009, he was elected a Foreign Member of the Polish Academy of Sciences. In 2012, he was elected a fellow of the Royal Society of Canada. In 2007, he received the Prestigious Norbert Wiener Award from the IEEE Systems, Man, and Cybernetics Society. He is a member of several editorial boards of international journals. He is a recipient of the IEEE Canada Computer Engineering Medal, the Cajastur Prize for Soft Computing from the European Centre for Soft Computing, the Killam Prize, and the Fuzzy Pioneer Award from the IEEE Computational Intelligence Society. He is vigorously involved in editorial activities. He is the Editor-in-Chief of *Information Sciences* and *WIREs Data Mining and Knowledge Discovery* (Wiley) and the Co-Editor-in-Chief of *International Journal of Granular Computing* (Springer) and *Journal of Data Information and Management* (Springer). He serves on an advisory board of IEEE TRANSACTIONS ON FUZZY SYSTEMS.

DARIUSZ CZERWINSKI (Member, IEEE) received the Ph.D. and D.Sc. (Habilitation) degrees from the Lublin University of Technology, in 2001 and 2014, respectively.

He was a Student with Kanazawa University, Japan. He is currently the Head of the Department of Computer Science, Faculty of Electrical Engineering and Computer Science, Lublin University of Technology. He is the author of more than 100 articles and the author/coauthor/editor of two books. His current research interests include fuzzy cognitive maps, artificial neural networks, the Internet of Things, cloud computing, and numerical analysis of physical phenomena. Some of his work has been on improving the understanding of big data analytics systems mainly through the application of data mining. He is a member of the Program Committee of international conferences, such as the Computer Networks Conference, the Conference on Computational Methods in Engineering Science, and the Conference on Computational Intelligence, Information Technology, and Systems Research.

KAMIL JONAK received the degree from the Faculty of Mechanical Engineering, Lublin University of Technology (LUT), and the Ph.D. and the D.Sc. (Habilitation) degrees in medicine from Medical University, Lublin, in 2019 and 2021, respectively. He works at the Department of Clinical Neuropsychiatry, Medical University of Lublin. He is currently the Head of the Department of Technical Computer Science, Faculty of Mathematics and Information Technology, LUT. His current research interests include biological neural networks, artificial neural networks, medical signals, image analysis, and the application of IT methods in healthcare.