



# Article A New Fuzzy-Based Classification Method for Use in Smart/Precision Medicine

Elena Zaitseva \*<sup>®</sup>, Vitaly Levashenko \*, Jan Rabcan and Miroslav Kvassay

Department of Informatics, Faculty of Management Science and Informatics, University of Zilina, 01026 Zilina, Slovakia; jan.rabcan@fri.uniza.sk (J.R.); miroslav.kvassay@fri.uniza.sk (M.K.) \* Correspondence: elena.zaitseva@fri.uniza.sk (E.Z.); vitaly.levashenko@fri.uniza.sk (V.L.)

Abstract: The development of information technology has had a significant impact on various areas of human activity, including medicine. It has led to the emergence of the phenomenon of Industry 4.0, which, in turn, led to the development of the concept of Medicine 4.0. Medicine 4.0, or smart medicine, can be considered as a structural association of such areas as AI-based medicine, telemedicine, and precision medicine. Each of these areas has its own characteristic data, along with the specifics of their processing and analysis. Nevertheless, at present, all these types of data must be processed simultaneously, in order to provide the most complete picture of the health of each individual patient. In this paper, after a brief analysis of the topic of medical data, a new classification method is proposed that allows the processing of the maximum number of data types. The specificity of this method is its use of a fuzzy classifier. The effectiveness of this method is confirmed by an analysis of the results from the classification of various types of data for medical applications and health problems. In this paper, as an illustration of the proposed method, a fuzzy decision tree has been used as the fuzzy classifier. The accuracy of the classification in terms of the proposed method, based on a fuzzy classifier, gives the best performance in comparison with crisp classifiers.

Keywords: fuzzy classification; fuzzy decision tree; data types; smart medicine; Medicine 4.0



Citation: Zaitseva, E.; Levashenko, V.; Rabcan, J.; Kvassay, M. A New Fuzzy-Based Classification Method for Use in Smart/Precision Medicine. *Bioengineering* 2023, *10*, 838. https://doi.org/10.3390/ bioengineering10070838

Academic Editors: Antonio Brunetti and Elena A. Jones

Received: 17 May 2023 Revised: 8 July 2023 Accepted: 13 July 2023 Published: 15 July 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

# 1. Introduction

There is a global consensus about the relevance of information technology (IT) for healthcare, within which it now plays an integral part [1–3]. Therefore, the creation of concepts or approaches that have a significant impact on IT also has a strong impact on medical applications. The first IT-based applications in medicine were associated with the use of decision-support systems and image processing, a document management system for medical records, and the development of evidence-based medicine [4]. The concepts of big data, the Internet of Things (IoT), cloud computing, artificial intelligence (AI), and other inventions have transformed modern technologies under Industry 4.0. The development of IT in recent times has resulted in such phenomena as "precision medicine", "smart medicine" or "Medicine 4.0" [3,5–7]. According to previous studies [3,5], an integral part of Medicine 4.0 or smart medicine is the use of AI-based applications in medicine (AI-based medicine), telemedicine, and precision medicine (Figure 1). AI-based medicine includes the most frequently used applications in healthcare, which allow the analysis of medical images and clinical and laboratory data. Artificial intelligence, including machine learning (ML) and big data processing, supports diagnosis and treatment in the healthcare environment [7–9]. Telemedicine using IoT will make seeing a patient or being seen by a clinician easier. Based on these technologies, such applications in the context of healthcare as electronic health records (EHR) [10], wearable devices (pulse oximeters or glucose monitoring) [5], and others have been developed. Precision medicine guides the course of treatment by using more comprehensive molecular diagnoses, such as genotyping or RNA expression.



Figure 1. The structure of Medicine 4.0.

The structure of Medicine 4.0, proposed by the authors of [3,5] (Figure 1), is created by data types, which are diverse in their nature and are obtained from various sources. These differences necessitate different types of data analysis, which, in turn, leads to different methods for processing. For example, the analysis of medical images requires methods to be in place for image processing and analysis; these methods cannot be those that are applied to expert data analysis and evaluation in the form of the physicians' notes drawn from EHR. The sharing of various IT methods, which are traditionally developed independently, is typical for medical applications. This problem can be considered partly from the point of view of the concept of big data. In their studies [2], the authors have shown that the principal impact on precision medicine has been the paradigm of big data. The role of big data analysis, in an environment where individual patient data, including their genes, environment, and lifestyle, are taken into account to identify the most suitable treatment. The methods for such data processing and analysis are mostly drawn from AI and ML.

AI in the context of health is used for a wide range of problems, from retrospective to prospective analysis. The methods and approaches of AI depend on many factors, such as the goal of the analysis (for example, data description, diagnostics, prediction, or prescription), data types (numeric, discrete, time factors, etc.), uncertainty (crisp or uncertain data, whether completely or incompletely specified) and other. The taxonomy of these problems, the methods, and the specifics of the data are illustrated in Figure 2 [11]. Four groups of applications, sorted according to this taxonomy, can be identified as relevant to healthcare, and precision medicine in particular, at the present time. The first of these groups can be identified as the collection of information from repositories, EHR, or biobanks [10]. The second and third groups can be interpreted as data analysis problems; the principal methods for their decision-making processes are clustering, classification, and regression (Figure 2), i.e., methods of ML. The initial data for these methods can be collected as data from biobanks and EHRs [12]. These methods enable the extraction of clinically meaningful insights from the data. Most medical applications currently deliver services within these three groups. In other words, these are effective methods and services for data collection, representation, preliminary analysis, and processing, in order to obtain clinically meaningful information for physicians. The fourth group of applications in healthcare is mostly present in the educational domain. The different simulators typically act as learning tools for use in medical universities [13,14].



Figure 2. Typical data types and ML methods for their processing and analysis.

As has been shown in several recent reviews [7,12,15,16], methods of classification, clustering, and regression are commonly used for medical applications (Figure 2). In some studies [17,18], association rules are considered to be one more alternative as a mathematical model for data analysis. They allow the identification of patterns, correlations, and causal patterns in the investigated samples. Among these methods, classification is the most frequently and successfully used technique. The greatest advances have been made in medical applications in which classification has been implemented based on neural networks or deep learning [19–21], decision trees or hierarchical analysis [22], and linear regression [23]. The most frequently used classification technologies in precision medicine are decision tree-based and neural network-based. The efficiency and advantages of the decision tree-based classifier lie in these factors: the classification result has good visualization and easy interpretation; a classifier can be inducted based on data with a small number of samples and/or a small number of attributes; there is no requirement to normalize the initial data. The undoubted advantages of neural networks are very good classification accuracy and the ability to process data with a very large number of samples. The only significant drawback of neural networks (deep learning) from the point of view of precision medicine is its representation in the form of a "black box", which does not allow researchers to retain at least some interpretability of the classification result. In some cases, to eliminate this shortcoming of neural networks, several classifiers are used simultaneously. There are studies in which an ensemble of classifiers is used. For example, in one study [24], random forests, an artificial neural network (ANN), and a support vector machine (SVM) were used for COVID-19 outbreak prediction. Therefore, despite the obvious advantages offered by neural networks, quite often, developers turn to such classifiers as decision trees [22,25,26], random forests [27,28], and fuzzy decision trees [29,30]. These classifiers offer good interpretability of the classification results, which is very important in precision medicine applications.

However, on the one hand, there are neural networks that offer good efficiency and applicability to big data, while on the other hand, there are decision tree-based classifiers that offer good interpretability of the classification result. The development of precision medicine applications should keep in mind the specifics of initial data for classifier induction. The data in precision medicine are of different types (heterogeneous). These data can be incomplete and uncertain, which is considered to be cognitive uncertainty [31]. Typically, this type of uncertainty is investigated using fuzzy logic [32,33]. Therefore, fuzzy classifiers can be considered an acceptable solution for classification problems in precision

medicine. Taking into consideration the interpretability of the result, this should be a fuzzy classifier based on a decision tree. At the same time, a new method for classifier induction should allow its development for the different types of initial data. In particular, we intend to consider and systemize classification signals [29,34], large-dimensional data [35,36], numeric data [37,38] and linguistic data [28,39]. An analysis of these studies shows that there are differences in the pre-processing steps of the initial data. A fuzzy classifier can be inducted, based on one method for the different types of initial data, if these data are transformed into fuzzy data at the pre-processing step.

In this paper, a new method for the classification of heterogeneous, incompletely specified, and uncertain data is proposed. The incompleteness and uncertainty of data are taken into account by the use of a fuzzy classifier because a data process based on fuzzy logic enables the achievement of the best result in the case of cognitive uncertainty. The heterogeneous initial data in the proposed method are processed by the special preliminary transformation of the initial data of different types into fuzzy data. This transformation is similar to data pre-processing for signal classification, based on a fuzzy classifier [29]. In the case of signal classification, the data pre-processing includes the feature extraction of the signal [40], feature selection [41], fuzzification [42,43], and classification. The large dimensional data can be transformed by feature selection and fuzzification methods before they are classified. Numeric or linguistic data need fuzzification via the appropriate methods [44,45]. Different fuzzy classifiers can be used in this method. However, in this paper, we consider one of the possible classifiers, namely, the fuzzy decision tree (FDT). This choice is motivated by the good interpretability and visibility of the results from FDT. Another advantage of this classifier is the possibility of using it to process data of both small and large dimensions. In this study, the various FDT applications for some medical problems are summarized. An analysis of the efficiency of classification based on fuzzy classifiers shows that these classifiers are effective and yield the best accuracy in terms of classification, in comparison with crisp classifiers. The proposed method also allows FDT induction regarding different types of data. The novelty of this method is the universalization of the classification, which improves the compatibility of the mutual analysis of classification results from different resources.

This paper is organized as follows. The specific medical data of selected types for use in the proposed method are discussed in Section 2. The structure and principal steps, along with the procedures for every step of the proposed method, are presented in Section 3. A detailed analysis of every introduced procedure is presented in Section 4. In Section 5, examples of the application of the proposed method with different data type classifications are discussed. In this section, the classifications of EEG signals (signals), laboratory data (large dimensional data), clinical data (numeric and categorical data), and expert data (fuzzy data) are presented. Finally, an analysis of the classification metrics for every problem is introduced to show the efficiency of the proposed method.

#### 2. Data Processing Specifics

## 2.1. Medical Data

The data used in precision medicine are big data. The typical characteristics of big data are volume, variety, and velocity (the "three Vs" of big data), which are then supplemented with such properties as value, veracity, and variability (the "six Vs" of big data) [46–48]. In big data processing, an emphasis is often placed on their large volume and the speed of occurrence (velocity). However, in the case of medical data, their heterogeneity is no less important and should be taken into account first [2,3,46]. The heterogeneity (variety) of medical data results from the use of a variety of available sources. Precision medicine uses comprehensive molecular diagnoses, which are based on the analysis of omics data (genotype, protein expression, RNA expression, etc.) [6]. Progress in IT and the emergence of the concept of Industry 4.0 have enabled the next step in medicine development, which is Healthcare 4.0. From the point of view of medical data analysis, the types of data involved in the processing are increasing [3], and the application of intelligence analysis methods is a

distinguishing characteristic of Healthcare 4.0. The sources of data in this domain might be sensors, images, gene arrays, laboratory tests, free text, demographic data, etc. (Figure 2). This means that different types of data are involved in the work of processing and analyzing their significance. As a rule, different types of data require different methods for their processing: the processing of omics data is different from the processing and analysis of clinical data or signals from sensors, and special methods are used for each of these types. Of course, there are fundamental approaches that are used for processing most of the different types of data (such as classification based on a neural network), but there are significant differences in the pre-processing of each data type, and in the interpretation of the results. This can be illustrated by the review of precision medicine data processing presented by the authors of [6]. In this context, a problem with the comparability and compatibility of the obtained results is revealed [49,50]. Decisions regarding the comparability and compatibility problem can cause the additional transformation of the result and may increase the computational complexity of data processing [51] or uncertainty regarding the eventual decision [52].

However, uncertainty not only arises as a result of the processing of medical data but is also inherent in the initial medical data. There are many factors that can cause uncertainty [53]. It can result from the poor quality of data measurement or evaluation, where the values of the initial data should be considered according to likelihood. Uncertainty can be caused by the incomplete specification of the data, where collected data contain missing values because some situations (samples) cannot be measured or evaluated. One more cause of uncertainty can be from data pre-processing, where, for example, a dimensional reduction can result in the loss of some information [54,55].

According to a previous study [31], uncertainty can be interpreted as being both statistical and cognitive. This statistical uncertainty is related to the randomness of nature and can be mathematically represented in terms of probability theory. Cognitive uncertainty is related to the level of uncertainty associated with cognition or the abstraction of reality [56]. Cognitive uncertainty includes such types of uncertainty as ambiguity and vagueness. Vagueness is the uncertainty due to "borderline situations", which is correlated with the difficulty of making exact decisions in borderline situations. Ambiguity is the uncertainty due to "alternatives" that exist in cases where the choice between alternatives is not exactly specified or some alternatives can be explained in a similar way. In many cases, the ambiguity and vagueness of medical data can be caused by experts' knowledge, which is typical in healthcare [53,56]. The incompleteness of such data can be expensive or even dangerous for patients. One typical approach for taking into account the uncertainty of different types of data is the representation of information as fuzzy data [57,58].

Therefore, the development of medical applications should take into account heterogeneous and uncertain medical data. In this paper, we propose a new method for medical data classification that allows the processing of different types of data, based on a single platform that can take into account the uncertainty and incompleteness of medical data. The specific feature of the proposed method is the use of the fuzzy-based processing of the data, which takes into consideration the uncertainty of the initial data for the resulting analysis.

# 2.2. Fuzzy Logic

Fuzzy logic introduces the notion of a membership degree [59]. A fuzzy set, A, with respect to a universe, U, is characterized by a *membership function*,  $\mu_A$ : U  $\rightarrow$  [0, 1], assigning an A-membership degree,  $\mu_A(u)$ , to each element of u in U.  $\mu_A(u)$  gives us an estimation of u that belongs to A. For  $u \in U$ ,  $\mu_A(u) = 1$ , which means that u is definitely a member of A, while  $\mu_A(u) = 0$  means that u is definitely not a member of A, and  $0 < \mu_A(u) < 1$  means that u is partially a member of A. If either  $\mu_A(u) = 0$  or  $\mu_A(u) = 1$  for all  $u \in U$ , A is a crisp set. In the opposite scenario, A is a fuzzy set. The cardinality measure of the fuzzy set A is defined by  $M(A) = \Sigma_{u \in U} \mu_A(u)$ , which is a measure of the size of A.

The degree of membership allows specific elements to be partial members of a fuzzy set. It presents a contrast with traditional set theory because a classical set is defined by crisp (exact) boundaries. Therefore, there is no uncertainty about the location of the set boundaries. A fuzzy set is defined by its ambiguous boundaries. Hence, there exists uncertainty about the location of the fuzzy set boundaries. Partial membership is represented by a membership degree that can be between 0 and 1, including both 0 and 1. Fuzzy logic provides very valuable flexibility in terms of reasoning, which makes it possible to take into account the various inaccuracies and uncertainties.

Fuzzy logic enables the processing of this uncertainty but, at the same time, provides for the use of special methods and algorithms for data analysis. In addition, the initial data should be transformed into fuzzy data. These conditions result in the development of new methods for medical data analysis, based on the use of fuzzy logic. Below is one of the possible methods for medical data analysis based on the classification methodology.

#### 3. A New Fuzzy-Based Method for the Classification of Medical Data

The methods used for medical data processing in precision medicine or in Medicine 4.0 should allow for the analysis of heterogeneous data (Figure 2). From the point of view of data analysis, the heterogeneity of the initial or source data is primarily associated with their processing. This process is most often defined as data pre-processing. Depending on the data type, the data can be transformed into another type, reduced, and cleared of noise and non-essential information. Depending on the procedures used in pre-processing, preliminary data can be interpreted as signals, images, high-dimensional data, numerical data, and expert data. Each type has its own pre-processing procedures. In this paper, the analysis of heterogeneous medical data according to classification is considered. A fuzzy classifier is used for the purposes of data analysis, to reduce the influence of pre-processing procedures on the classification result. The merging of different types of data is carried out by converting them to fuzzy data.

The core of the proposed method is a fuzzy classifier (Figure 3). As was shown in previous studies [29,30,32,33], the use of a fuzzy classifier makes it possible to increase the accuracy and efficiency of the analysis of incompletely specified and uncertain data [30,32] or signals [29]. In previous studies [54,55], the authors have shown that the pre-processing of a signal for its classification leads to the loss of some information. This loss is not large, but it can have an influence on the classification efficiency. The transformation of the pre-processing result into the fuzzy domain extends the classification attribute value borders and decreases the impact of lost information. In other studies [29,37] the efficiency of fuzzy classifiers for use with large dimensional data has been investigated.

The typical data pre-processing seen in signal classification includes the procedures of feature extraction and feature selection (dimension reduction). The procedure regarding feature extraction enables the discovery of information that can effectively reflect the specifics of the analyzed signal and identify its classification. The pre-processing of large dimensional data only needs to achieve dimension reduction, which can be performed via a procedure similar to signal classification. A dimension reduction procedure deletes unnecessary or redundant features by evaluating their usefulness and impact on classification efficiency. Feature selection and, as a result, the reduction in dimension of classified data, are required to simplify data visualization and comprehension and to minimize the time needed for classifier induction by decreasing the training time. The use of fuzzy classifiers determines that the input data for the classification should be fuzzy. Therefore, other types of input data (attributes) should be transformed into fuzzy data via fuzzification. The transformation enables the introduction of fuzzification to the proposed method. This procedure transforms each numeric attribute value into a fuzzy value with a membership function and is used for numeric data to achieve the unification of the utilized classifier. The expert (linguistic) data can be classified without pre-processing.



**Figure 3.** The structure of a new method for the classification of heterogeneous and uncertain medical data.

The structure diagram in Figure 3 shows that the proposed method enables the classification of data of various types, such as signals, high-dimensional numerical arrays, numerical data, and expert data. All these types of data are converted to fuzzy data using sequentially applied pre-processing procedures. The classifier is developed according to specified machine-learning algorithms for every type of data. The initial data for a classifier development create a table (dataset) that represents the fuzzy values of the output target attributes, depending on the fuzzy values of input and independent attributes. This table does not include all the possible combinations of values for the input attributes. Therefore, this table is not a decision table and cannot be used for the representation of complete knowledge about this object or process. The developed classifier permits the creation of a decision table for the investigated object or process, for analysis, or for use in the classification of a new sample.

Each of the pre-processing procedures can be implemented, based on various algorithms and methods. It should be noted that the modification of these procedures can have a significant impact on the classification result. Unfortunately, the choice of the most efficient implementation method for each procedure is determined experimentally for each studied dataset [60–62]. There follows a brief overview of the most commonly used procedures of pre-processing.

# 4. Principal Steps of the New Method

# 4.1. Feature Extraction

Feature extraction is the first procedure in the pre-processing of a sensor's dataset or signal [40]. The specific features of the signal should be extracted at the first step of signal pre-processing. The major goal of feature extraction is to obtain information that can effectively reflect the specifics of the analyzed signal and identify its class. Moreover, sensory data or signals are represented in the time spectrum, which is not suitable for classification without additional transformation. The accuracy of classification is mostly determined by the extracted features of the signal.

The feature extraction of sensor data or a signal is mostly implemented based on the methods in three groups: time, frequency, and time-frequency domains [63,64]. Some researchers consider a further group, namely, nonlinear feature analyses [65]. The methods of transformation of a signal in the time-frequency domain are the most commonly applied techniques for signal feature extraction.

Time-domain features are computed based on the signal amplitude according to time. Therefore, these methods are known as amplitude-based methods for feature extraction. The advantage of these methods is their low computational complexity. Moreover, this type of analysis often does not require additional signal transformation. There are many techniques used in the time domain for signal feature extraction. These methods include the histogram analysis method [66], higher-order crossing [67], and principal component analysis (PCA) [68]. Amplitude analysis can also be based on statistical features [69].

Features extraction based on Fourier transform methods [70] (fast Fourier transform and discrete Fourier transform) is conducted in the frequency domain (Table 1). Fourierbased and other frequency-based domain analysis methods convert time-domain signals to frequency-domain signals, in order to extract the frequency-domain features. These transforms are often used for EEG signal analysis and classification [70–72]. In addition to Fourier transform, the group of frequency-based feature extraction methods includes various techniques for the analysis of power spectral density (PSD) [71] and differential entropy (DE) [72].

**Table 1.** Comparison of the efficiency of the proposed method based on FDT and other classifiers for signal classification.

<b>Classification Algorithm</b>	Туре	Accuracy	Specificity	Sensitivity	Precision	F1 Score
FDT	Fuzzy	0.995	0.993	0.996	0.981	0.989
Decision Tree, C4.5	Non-fuzzy	0.981	0.969	0.992	0.978	0.984
Fuzzy Naive Bayes classifier	Fuzzy	0.952	0.917	0.999	0.862	0.924
Naive Bayes	Non-fuzzy	0.962	0.928	0.987	0.949	0.968
Fuzzy Multi-Layer Perceptron	Fuzzy	0.948	0.889	0.997	0.919	0.956
Multi-Layer Perceptron	Non-fuzzy	0.942	0.881	0.991	0.913	0.950

The time- and frequency-based analysis methods used here are effective for linear and quasi-stationary signals [64]. In healthcare, signals are not linear and are quasi-stationary, for example, EEG signals. Time-frequency domain analysis methods (spectral methods) combine information from the time and frequency domains and permit time-frequency domain-localized analysis. The most effective and widely used methods of analysis in this group are wavelet transform [73] and Welch's method [74]. Short-time Fourier transform (STFT) [75] and the Hilbert Huang transform (HHT) [76] are also essential time-frequency domain analysis methods.

There are studies in which new methods or combinations of previously known methods are developed for effective feature extraction of the signal. A new method for feature extraction for EEG signals, based on a combination of Fourier and Wavelet transforms using fuzzy entropy, is developed in an earlier study [77]. The authors of [78] propose a new time-frequency transformation by using joint approximate diagonalization for feature extraction.

A very brief overview of the feature extraction of sensory data or signal methods is presented below. Here, it should be noted that the determination and development of the most effective feature extraction procedure is carried out experimentally for each type of signal. The feature extraction procedure is very important in terms of data pre-processing. However, the data obtained after this procedure have a large dimension, which must be reduced using the following procedure to ensure efficient classification.

#### 4.2. Dimensionality Reduction

Dimensionality reduction or feature selection represents the next step in data preprocessing. This procedure can be used after the process of feature extraction or for processing initial data of larger dimensions. Feature selection is a procedure that is used to delete unnecessary (or redundant) features by evaluating their usefulness and their impact on classification efficiency in order to obtain the best results, based on the least possible amount of processed data [41]. Feature selection and, as a result, a reduction in the dimension of classified data are required to simplify data visualization and comprehension and to minimize the time needed for classifier induction by decreasing the training time.

The most frequently applied procedures for feature selection are based on the techniques of linear discriminant analysis (LDA) [79], principal component analysis (PCA) [80], and independent component analysis (ICA) [81]. These techniques can be used to map a high-dimensional input vector into a low-dimensional vector. This transformation changes the physical context and nature of the initial features, which is not appropriate for some classification problems. However, these techniques are effective for signal classification [54]. PCA represents a more universal method for signal classification [79]. The use of PCA transforms the initial feature matrix with *n* features,  $Y = (Y_1, \ldots, Y_s)$ , into a new matrix,  $X = (X_1, \ldots, X_n)$ , in which the features (columns) are known as "principal components" and  $s \ge n$  [79]. In order to reduce the dimensions of the matrix, we must select some of the most important principal components. The importance of the principal components can be defined using component variance. The variance of the components indicates variability in the data.

There are techniques for feature selection that do not transform the initial feature vector [82,83]. Feature selection methods permit the exclusion of secondary features and choosing those that will have the maximum impact on the classification accuracy. These methods can be considered as three groups [82,83]: the filter, wrapper, and embedded methods.

Filter-based methods use ranking to select the relevant features. The relevance is evaluated based on mutual information, the Pearson correlation coefficient, or other criteria [84]. The computational complexity of these methods is lower, and they can be used for any classifier induction. Each feature is evaluated separately using filter methods; therefore, some features can be removed according to an individual evaluation but they may become relevant when combined with others, or features that are considered individually to be relevant may result in unnecessary redundancies. Wrapper algorithms [85] execute feature selection in the context of (and in conjunction with) the classification algorithm. These methods can be interpreted as an optimization algorithm that employs the classification results as the target function. The computational complexity of wrapper methods is high. Finally, embedded methods perform attribute selection in the process of classifier induction and depend on the classifier and algorithm of this induction [82].

#### 4.3. Fuzzification

Fuzzification is the procedure for data pre-processing used in the proposed method, which transforms linguistic data [42,44] or crisp numeric data [43,45] into fuzzy data (fuzzy attributes). Fuzzy attributes have values that are defined using the membership function (see Section 2.2). Every fuzzy attribute,  $A_i$ , consists of the defined number  $m_i$  ( $m_i \ge 2$ ) in linguistic terms. The *j*-th linguistic term of  $A_i$  is defined by the fuzzy set  $A_{i,j}$  ( $j = 1, ..., m_i$ ). Many methods are used for the fuzzification of both types of data. One of the first methods used was proposed by Zadeh [59], based on the use of non-fuzzy cardinality. Popular methods for the fuzzification of linguistic data were first developed by Yager [56,86]. A new definition for a linguistic quantifier was introduced elsewhere [87], while a two-step method for fuzzification and fuzzy quantification was introduced by the author of [44]. A

review and analysis of many methods for the fuzzification of linguistic data are laid out in Ref. [42].

There are several fuzzification methods for crisp data. In this case, fuzzification converts crisp numeric attributes (principal components) that have been obtained after dimensionality reduction into fuzzy attributes  $A_i$  (i = 1, ..., n). The numeric attribute  $X_i$  is defined by a vector of real scalar values ( $x_1, x_2, ..., x_k, ..., x_K$ ), where K defines the number of samples. The definition of this fuzzy set,  $A_{i,j}$ , with respect to  $X_i$  is created by a membership function,  $\mu_{A_{i,j}}(x)$ :  $X_i \to \{0, 1\}$ , which provides a membership degree  $\mu_{A_{i,j}}(x)$  for each x ( $x \in X_i$ ). The membership degree  $\mu_{A_{i,j}}(x)$  determines how strongly the element, x, is a member of the fuzzy set,  $A_{i,j}$ . Formally speaking, fuzzy set  $A_{i,j}$  is defined as an ordered set of pairs,  $A_{i,j} = \{(x, \mu_{A_{i,j}}(x)), x \in X_i\}$ , where:

- (a)  $\mu_{A_{i,i}}(x) = 0$ , if, and only if, x is not the member of set  $A_{i,j}$ ;
- (b)  $0 < \mu_{A_{i,i}}(x) < 1$ , if, and only if, *x* is not the full member of set  $A_{i,j}$ ;
- (c)  $\mu_{A_{i,i}}(x) = 1$ , if, and only if, x is the full member of set  $A_{i,j}$ .

Fuzzification of crisp data can be implemented based on *g*-fuzzification [88], an evaluation of the histogram of variables [89], entropy analysis [90], and clustering [43,91,92]. Well-known clustering-based methods for fuzzification include the K-means algorithm and the fuzzy C-means method (FCM).

#### 4.4. Fuzzy Classification

There are many different fuzzy classifiers, for example, the fuzzy naïve Bayes classifier [93], fuzzy classification rules according to the algorithm reported in Ref. [94], neural networks [95], and decision tree-based classifiers [27–29,96]. The choice of one of these classifiers is implemented by an experimental evaluation of result efficiency (including accuracy, sensitivity, availability, and other metrics [62,97]). The properties of the classifier can also have an impact on the choice. For example, neural networks are effective for large dimensional data and are not suitable for data from small samples. However, Bayes classifier or tree-based methods are effective for such data. Note that the influence of the procedures of pre-processing is possible regarding classification efficiency. Therefore, the evaluation of a specified classifier's efficiency should be implemented for different procedures of pre-processing. In this paper, we propose a technique using a decision tree-based classifier, specifically, FDT [29,30]. This classifier has very good interpretability, which is important for the medical domain.

There are many different methods for FDT induction [96,98]. In this study, FDT induction is implemented based on cumulative information estimation (CIE), which was introduced by the authors of [99]. CIE evaluates the degree of influence of the input attribute on the resulting (output) attribute, in the context of information theory. In contrast to an informational assessment, which takes into consideration the average values of each input attribute, the CIE technique allows us to take into consideration the values of all observations and the measurements of each input attribute. The structure of FDT (the sequence and correlation of nodes) is determined, based on the CIM. Additionally, when building a tree, the parameters a and b are used. The values for the pruning parameters  $\alpha$ and  $\beta$  affect the size and efficiency of the inducted FDT [98]. The decrease in parameter  $\alpha$  and the increase in parameter  $\beta$  result in an FDT extension (due to increasing the numbers of levels and nodes) and a decrease in the classification accuracy because of the redundant information (noise) involved in FDT. An increase in parameter  $\alpha$  and a decrease in parameter  $\beta$  lead to a reduced FDT size and a decrease in classification accuracy because essential information can be lost. The calculation of CIM and the process of choosing pruning parameters  $\alpha$  and  $\beta$  is discussed in Ref. [98] in detail.

A decision tree is a flowchart-like structure in which each internal node represents a check on an input attribute, each branch represents the outcome of this check, and each leaf node represents a value of the output attribute (class label).

The fuzzy classification efficiencies are evaluated based on several metrics, which are computed according to the results of the following classifications (classification instance) [62,97]: true positive (TP), true negative (TN), false positive (FP), or false negative (FN). The metrics of this classification are calculated, based on the classification instance:

• Accuracy is the ratio of correctly predicted instances to all predicted ones:

$$Acc = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$
(1)

• Specificity is the proportion of correctly identified of true negative results:

$$Spec = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}} \tag{2}$$

• Sensitivity is the probability of a correct initial prognosis:

$$Sens = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{3}$$

 Precision is the proportion of correct positive predicted instances to the total number of instances in terms of positive predicted instances:

$$Prec = \frac{TP}{TP + FP}$$
(4)

• The F1 score is the harmonic mean of sensitivity and precision:

$$Fsc = \frac{2TP}{2TP + FP + FN}$$
(5)

# 5. Examples of Method Application

In this section, illustrative examples of the proposed method are introduced. Signal classification, based on the proposed method, is shown in the context of EEG signal classification to detect epilepsy seizures [29]. Laboratory research into tumors, based on an analysis of blood parameter classification, is considered an example of large dimensional data classification [35]. Numeric data classification is illustrated by the timing of tracheostomies in COVID-19 patients [37]. Expert data classification is considered for an analysis of the influence of the human factor in a medical team [30]. In this paper, we do not induct the FDTs for the classification of the EEG signal, the timing of tracheostomy in COVID-19 patients, and human factor influence in a medical team. Instead, we show how these FDTs were inducted in previous studies [29,30,37] to illustrate the application of the proposed method (Figure 3).

The initial data for the considered classifiers' induction have been introduced in our previous studies on signal classification [29], large dimensional data [35], numeric data [37], and expert data [30]. In these studies, we have discussed the induction of FDTs in detail. In each of the studies on FDT induction, the calculation of CIM and the pruning parameters  $\alpha$  and  $\beta$  are discussed in detail. The comparison of inducted FDTs with other classifiers is also considered. In this section, the results of previous investigations and comparisons are used to illustrate the proposed new method.

The inducted classifiers used in this study are evaluated using Equations (1)–(5) for the creation of a machine learning model on future (unseen/out-of-sample) data. This metric calculation is based on the application of the hold-out method. The initial dataset is divided into 2 samples; then, we build a model with the first (training) sample. The second (testing) sample is used for verification of the built model. Usually, the proportion of the dataset division is 70:30. According to the second method in the evaluation of the classifiers' efficiency, random subsampling, the previous hold-out method is applied many times to achieve different divisions of the initial dataset.

## 5.1. Signal Classification

The EEG signal has many applications in medicine for diagnosing disease [63,65,67,68, 71,72]. Often, these signals are used for neurology diagnostics in epilepsy patients [29,68,78]. There are some public datasets to study these signal classifications, which include records (samples) of EEG signals. In a recent study [29], the dataset from [100] has been used for the FDT induction and classification of epileptic seizures. This dataset includes 500 samples, comprising 23.6-second signal records. The dataset consists of 5 subsets of 100 samples. These subsets are collected for different conditions and patient states and are indicated by the letters A, B, C, D, and E. Subsets A and B were collected to include those persons who do not suffer from epilepsy, in different conditions. These conditions comprise having the eyes open or closed because this affects the electrical activity of the brain. The difference between these subsets is that for subset A, the patients had their eyes open, while for subset B, their eyes were closed. The subsets C, D, and E were collected from persons who suffered from epilepsy. The signals in subset E were collected for patients at the time of a seizure. The subsets C and D were collected during seizure-free intervals. The difference between subsets C and D is the zones of the EEG signal recording: samples of subsets D and C were recorded from within the epileptogenic zone and the hippocampal formation of the opposite hemisphere of the brain, accordingly. The problem of EEG signal classification can be considered as a classification into two classes, as ((A, B) and (C, D, E)) or ((A, B, C, D) and E)). In the same study [Rab2022], a second type of classification has been studied: (A, B, C, D) and E.

The proposed method for this case includes all the procedures of pre-processing (Figure 3). The procedures of fast Fourier transform (FFT) are used for feature extraction, those of PCA for dimensional reduction, and those of FCM for fuzzification. The initial data from EEG signal records in this study are cut and collected in a set of 4000 2.95-second records. This set of 4000 records of EEG signals are analyzed using FFT. The Fourier specter transforms to 8 attributes via PCA. The number of principal components has been obtained, based on an experimental investigation: the decrease in the number of principal components decreases the efficiency of the classification process, while an increase in the number of principal components increases the computational complexity of classification. This experiment is discussed in more detail in an earlier paper [29]. The procedure of FCM permits the representation of eight principal components by eight fuzzy attributes.

In Table 1, an evaluation of the inducted FDT for EEG signal classification, in comparison with other classifiers, is shown. The comparison of fuzzy and non-fuzzy classifiers shows that the classification accuracy of the fuzzy classifiers yields the best performance in comparison with the similar non-fuzzy classifier. An analysis of the metrics for FDT shows that this classifier offers the best results in terms of accuracy, sensitivity, and F1 score, along with the second-best results in terms of precision and specificity. According to the same study [29], the inducted FDT is the best classifier for the identification of epilepsy seizures, based on the EEG signal.

The classification of the EEG signal is effectively implemented, based on fuzzy classifiers, which confirms the effective application of the proposed method for signal classification.

# 5.2. Large Dimensional Data Classification

Data are considered highly dimensional if the number of attributes is larger than the number of instances. High-dimensional datasets are common in the biological sciences. Fields of study such as genomics and the medical sciences will often use both tall and wide datasets that can be difficult to analyze or visualize using standard tools. An example of high-dimensional data in the biological sciences may include data collected from hospital patients, recording their symptoms, blood test results, behavior, and general health, resulting in datasets with large numbers of features.

Future experimental studies should be implemented for large dimensional data. Scientific laboratory research into the diagnosis of a tumor based on an analysis of blood parameters can be considered as an example. This laboratory study was performed on rats; data were created after monitoring Sprague Dawley rats [35]. The animals were adapted to standard vivarium conditions (temperature, humidity, a regular light and dark regimen). Data included 78 instances (rats). Monitoring was based on rat blood analysis. Blood from all the experimental animals was collected at a single point and then analyzed. After blood analysis, we obtained 186 input attribute values for each rat. These rats were then divided into two groups: the control and tumor groups. The goal of the prediction was to find out if the animal was suffering from a tumor.

The analysis of this data, according to the proposed new method, includes preprocessing and classification based on a fuzzy classifier, which, in this case, is FDT (Figure 3). Here, the pre-processing consists of the procedures of dimensional reduction and fuzzification, PCA, and FCM. The initial data are transformed into a new set of principal components, which consists of 5 input attributes.

We made a comparison of the classification results of the proposed method and other non-fuzzy classifiers for the same initial data, using the decision tree, naïve Bayes, neural network, k-nearest neighbor, and support vector machine (SVM) techniques. The experiments were implemented in MATLAB. For each classifier, we ran a procedure to find the best combination of input parameters. The procedure was based on multiple runs of the classifier using different values for the input parameters. The best classification result identified is shown in Table 2.

**Table 2.** Comparison of the efficiency of the proposed method, based on FDT and other classifiers for large dimensional data.

Algorithm	Туре	Accuracy	Specificity	Sensitivity	Precision	F1 Score
FDT	Fuzzy	0.872	0.892	0.854	0.897	0.875
Decision Tree, C 4.5	Non-fuzzy	0.859	0.892	0.829	0.895	0.861
Naïve Bayes	Non-fuzzy	0.833	0.757	0.902	0.804	0.851
Neural Network	Non-fuzzy	0.859	0.838	0.878	0.857	0.867
K-nearest Neighbor	Non-fuzzy	0.795	0.811	0.780	0.821	0.800
SVM	Non-fuzzy	0.846	0.784	0.902	0.822	0.860

The classification of the data from the laboratory research shows the best result, according to the considered metric for the fuzzy classifier, which is FDT. This study shows that the proposed method is effective for large dimensional data and the use of a fuzzy classifier is effective.

#### 5.3. Numeric Data Classification

In an earlier study [37], an analysis of the timing of tracheostomy in COVID-19 patients was based on the classification of initial data via FDT. The analysis, which was based on FDT, was developed in an investigation by the authors of [100], and included the table of supplementary materials (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7673767/, accessed on 17 November 2020). The data were collected from 177 health records from anonymous COVID-19 patients in the ICU at Guy's and St Thomas' National Health Service (NHS) Foundation Trust. The initial data of the investigated dataset comprise 29 input attributes for two groups. The attributes of the first group were obtained as the baseline characteristics, while the attributes of the second group were obtained during 14 days of hospitalization. The analysis was focused on an evaluation of tracheostomy timing and the prediction of a patient's survival.

Baseline dataset characteristics can be interpreted as input attributes that are obtained immediately after the admission of a patient or by an analysis of the patient's health records. The attributes of the first part, which can be obtained immediately, are age, gender, ethnicity, body mass index (BMI), and the acute physiology and chronic health evaluation II (APACHE II) score. The baseline characteristics, drawn from the patient's health records, may be the confirmation of diabetes mellitus, hypertension, ischemic heart disease, chronic obstructive pulmonary disease, asthma, and chronic kidney disease. One more input attribute of the baseline characteristics is thromboembolism (pulmonary, venous, or multiple). These attributes, which were identified based on the baseline characteristics, are categorical (except for age and BMI, which are numerical).

The second group of input attributes was collected from the patient's vital signs, markers of acute respiratory failure, and serum-based biomarkers for the severity of the disease. These attributes were measured repeatedly at different time points (after 24 h of clinical care on days 7, 10, and 14). If a patient died or was disconnected from mechanical ventilation, the last measured values were used. These attributes are identified as PEEP (positive end-expiratory pressure), FiO2 (fraction of inspired oxygen), PaO2 (partial pressure of oxygen), the PF ratio, CRP (C-reactive protein), ferritin, D-dimer, temperature, vasopressors, RRT (renal replacement therapy), and ECMO (extracorporeal membrane oxygenation). The attributes of the second group are numerical.

The input attributes for the considered problem are both categorical and numerical. The categorical data regarding patients can contain ambiguity, for example, due to the problem of the unambiguous interpretability of information in health records. The measured numerical data can have vagueness, which, for example, can be caused by a measuring device error. Therefore, fuzzy classification is more acceptable for the considered problem of facilitating decision-making as to the optimal timing of tracheostomy for prolonged respiratory weaning in critically ill COVID-19 patients.

The evaluation of different classifiers is shown in Table 3. In this study, for numerical data, the best decision is obtained with the fuzzy classifier (FDT). Therefore, we can see that the proposed method is also effective for numeric data analysis.

Classification Algorithm	Туре	Accuracy	Specificity	Sensitivity	Precision	F1 Score
FDT	Fuzzy	0.867	0.563	0.921	0.921	0.921
Decision Tree, C4.5	Non-fuzzy	0.835	0.357	0.91	0.9	0.905
Fuzzy Naive Bayes classifier	Fuzzy	0.816	0.143	0.921	0.872	0.896
Naive Bayes	Non-fuzzy	0.816	0.143	0.921	0.872	0.896
Fuzzy Multi-Layer Perceptron Multi-Layer Perceptron	Fuzzy Non-fuzzy	0.857 0.845	0.563 0.357	0.91 0.921	0.92 0.901	0.915 0.911

**Table 3.** Comparison of the efficiency of the proposed method, based on FDT and other classifiers of numeric data.

# 5.4. Expert Data Classification

The last example provided for the proposed method considered the problem of the development of a mathematical model for the human reliability analysis of a medical team. The analyzed team included a doctor and two nurses. The goal of this problem is to evaluate the impact of human (medical) error on patient safety [30] (see Table 8 in Section 7). The mathematical model for human reliability analysis should classify all possible performance levels of the team's members into three classes: a fatal medical error (set at 0), some imperfection in the patient's care (set at 1), and patient care without any complications (set at 2). Therefore, the output attribute B has three values:  $B_0$ ,  $B_1$ , and  $B_2$ . The number of input attributes is defined according to the number of team members, which is 3. Two attributes,  $A_1$  and  $A_2$ , are interpreted here regarding the nursing performance and have only two possible states: error (state 0) and error absent (state 1). The doctor's performance is considered attribute  $A_3$  and has four levels: from a fatal doctor error (set at 0) to the doctor's perfect work (set at 3). The impact of the mistakes made by a doctor or nurse on

patient safety must be indicated and evaluated, based on a method of human reliability evaluation.

The initial data comprise expert evaluations of all 10 samples [30]. These data are incompletely specified because the sum of all possible situations is 16. Every value for the performance of each team member is indicated by a certainty. These data can be considered fuzzy data; thus, pre-processing is not needed for fuzzy classifier induction in this case. Therefore, the FDT can be inducted based on the initial data, without the need for additional transformation. An evaluation and comparison with other classifiers are presented in Table 4.

**Table 4.** Comparison of the efficiency of the proposed method, based on FDT and other classifiers, for expert data.

<b>Classification Algorithm</b>	Туре	Accuracy	Specificity	Sensitivity	Precision	F1 Score
FDT	Fuzzy	0.934	0.456	0.823	0.893	0.944
Fuzzy Naive Bayes classifiers	Fuzzy	0.915	0.372	0.798	0.852	0.902

The definition of the initial data values with certainty permits their interpretation as fuzzy data. The fuzzy classifier is inducted without the pre-processing of initial data, but a crisp classifier induction needs the initial data de-fuzzification. This procedure causes the classification accuracy to decrease. Only fuzzy classifiers should be considered for the comparison. There are specifics to this data that comprise the small dimension of the sample and a small number of input attributes. For such data, the use of a neural network as a classifier is impossible. A neural network cannot be trained on such a small dataset. Therefore, one classifier was inducted for the comparison with FDT, which is a fuzzy naive Bayes classifier. A comparison of the metrics of the classification efficiency in Table 4 shows that the FDT is best suited for the considered problem.

# 6. Discussion

In this paper, the specifics of the classification of various data are considered and a new classification method is proposed to classify certain types of medical data. In particular, the proposed method (Figure 3) permits the classification of signals (sensors data), large dimensional numerical data, numerical, linguistic, and categorical data, and fuzzy data. Depending on the data type, the number of successively used pre-processing procedures changes. Signal pre-processing includes feature extraction, feature selection or dimensionality reduction, and fuzzification. Accordingly, for large dimensional data, the procedures for feature selection or dimension reduction and fuzzification are applied. Pre-processing consists of only one fuzzification procedure for numerical, categorical, and linguistic data. The procedure of fuzzification is used in pre-processing for all data types. It is due to the presence of a fuzzy classifier in the proposed method. As shown in previous studies [29,30,37], fuzzy classifier exploitation can improve the efficiency of classification. It is caused by the loss of some useful information in the relevant step of data pre-processing; first of all, in the procedure of dimensional reduction or feature selection [54,55]. Transformation of the input attributes before the classification from crisp to fuzzy allows us to "smear" the crisp values and thereby expand their coverage range, leading to improved classification accuracy. It is necessary to note that the different procedures of pre-processing impact the classification result. Unfortunately, there is no way to predict in advance the effectiveness of a particular procedure for data pre-processing [60]. Therefore, for every one of the procedures used in pre-processing, a short review of possible methods for its implementation is introduced. In this study, the procedures of FFT, PCA, and FCM are used for feature extraction, dimensional reduction, and fuzzification, respectively.

The classification process for all the considered problems has been implemented, based on FDT. This classifier can be effective for processing data with numerous input attributes and data with some input attributes. It can also be used for data comprising very

small samples (see Section 5.4) without additional transformation, and also offers good interpretability and visibility. Of course, other classifiers can also be used instead of FDT. Similar to pre-processing procedures, the performance and efficiency of any classifier for each particular case cannot be predicted [60]. Using FDT for the considered problems in this study yields acceptable efficiency.

In this paper, the classification problems of the EEG signal (Section 5.1), scientific laboratory research on the diagnosis of a tumor, based on an analysis of blood parameters (Section 5.2), the timing of tracheostomy in COVID-19 patients (Section 5.3), and the expert data analysis of possible medical errors in a team (Section 5.4) have been considered. The data in these problems are different: signals, large dimensional data, numeric and categorical data, and expert (fuzzy) data. The proposed method (Figure 3) has been used in all these problems. The comparison of classification Equations (1)–(5) for FDT and other classifiers (first, for crisp classifiers) shows the efficiency of the proposed method.

For example, Table 1 (EEG signal classification), Table 2 (the diagnosis of a tumor based on an analysis of blood parameters), and Table 3 (timing of tracheostomy in COVID-19 patients) introduce the metrics of classification efficiency for several classifiers. The classification efficiencies are evaluated for the different data types: signal data (Table 1), numeric large dimensional data (Table 2), and numeric data of limited dimensionality (Table 3). The comparison of FDT and the decision tree inducted using C 4.5 methods show that these metrics are somewhat better suited for FDT for all data types. In one study [55], it has been shown that signal preprocessing before classification can cause a loss of useful information and, according to a more recent study [Rab2022], the transformation of preprocessing results in a fuzzy domain that permits the leveling out of this loss by expanding the classified value boundaries. Fuzzy classifiers have similar or better efficiencies regarding classification in comparison with crisp classifiers of all types: fuzzy naive Bayes classifiers for signal classification (Table 1) and numeric data (Table 3) have the best metrics for classification than crisp versions of the naive Bayes classifiers. A similar result is achieved in a comparison of the fuzzy multi-layer perceptron and the crisp multi-layer perceptron.

Among the considered examples, the one that depicts the analysis of medical errors using a human reliability analysis method should be noted. The initial expert data are indicated in the form of values with certainty, allowing us to interpret them as fuzzy data. Therefore, for this data, a fuzzy classifier should be used (Table 4). In this problem, the sample under study is very small (only 10 samples). Therefore, the choice of classifiers is very limited. In particular, neural networks for such a small sample cannot be used. FDT and the fuzzy naive Bayes technique have been inducted for this problem. The considered metrics indicate that the best efficiency was achieved by FDT (Table 4) and, in addition, it offers good interpretability of the result. This classifier was used in an earlier study [30] for the development of a mathematical model for a reliability analysis of the human factor in medical systems.

Therefore, the considered examples and the previous studies [29,30,37] show that our proposed method for the classification of different data types in smart and precision medicine, based on fuzzy classifiers, is efficient.

# 7. Conclusions

The most frequently considered problem in medical applications and services is result prediction, which depends on new initial data based on decisions that are already known. This problem is typical for both AI-based medicine and precision medicine (Figure 1). In fact, this can be considered a classification problem. The specifics of modern medical applications and services demand that such a prediction or classification must be performed for different types of data, and, as a rule, the classification of different types of data is implemented using different methods and algorithms. Typically, differences in these methods are found in the pre-processing procedures regarding the initial data (input attributes).

In this paper, a new method for the classification of different types of data is proposed; in particular, the classification of signal, large dimensional, numeric, or expert (linguistic or fuzzy) data is shown to be possible. One of this method's advantages is the use of well-known procedures for data pre-processing: feature extraction, feature selection or dimensional reduction, and fuzzification. This permits resources for the method development to be decreased. The novelty of this method is the fuzzy classifier technique used, along with the new interpretation and exploitation of known methods and algorithms to address the new problem of the classification of different data types based on a single classifier. The unification of the classifier will make it easier to combine and jointly analyze the results from data of various types. According to the experimental evaluations, fuzzy classifiers offer better classification efficiency for the considered data types. In this paper, FDT is applied as a fuzzy classifier because it offers good interpretability of the results in comparison with other classifiers. However, other fuzzy classifiers can also be used. In future investigations, the influence of different methods and algorithms regarding the procedures of data preprocessing will be investigated. The influence of a membership function on method efficiency will also be considered.

In this paper, the proposed method was used to classify each data type separately. This suggests that additional procedures will be required to analyze data of different types that are mixed together. This is a disadvantage to the proposed method. One final restriction of this method is discrete data processing. Continuous data should be discretized before the classification proceeds, based on the proposed method. In further studies, the different data types under consideration will be combined into a single sample, not in order to consider the processing of each of the presented data types independently, but instead to provide for their joint processing. In further studies, the proposed method will be supplemented with other data types that are widely used in medicine. One of the first additional types of data to be considered will be medical images.

**Author Contributions:** Conceptualization, E.Z.; methodology, E.Z. and V.L.; software, J.R.; validation, M.K.; formal analysis, E.Z. and V.L.; investigation, E.Z. and V.L.; resources, E.Z. and M.K.; data curation, J.R.; writing—original draft preparation, E.Z. and V.L.; writing—review and editing, E.Z.; visualization, V.L.; supervision, E.Z.; project administration, V.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work has been supported by the grant APVV 18-0027, "New method development for the reliability analysis of complex systems" of the Slovak Research and Development Agency, and grant VEGA 1/0165/21 "New approaches for the reliability analysis of non-coherent systems" of the Ministry of Education, Science, Research, and Sport of the Slovak Republic, and the Integrated Infrastructure Operational Program for the project "Integrative strategy in the development of the personalized medicine of selected malignant tumors, and its impact on quality of life", IMTS: 313011V446, co-financed by the European Regional Development Fund.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data for the experimental study in this paper are from previously published investigations [29,30,35,37] (see Section 5).

Conflicts of Interest: The authors declare no conflict of interest.

# References

- Wang, W.; Krishnan, E. Big Data and Clinicians: A Review on the State of the Science. JMIR Public Health Surveill. 2014, 2, e1. [CrossRef] [PubMed]
- Iwashyna, T.J.; Liu, V. What's So Different about Big Data?. A Primer for Clinicians Trained to Think Epidemiologically. Ann. Am. Thorac. Soc. 2014, 11, 1130–1135. [CrossRef]
- 3. Haleem, A.; Javaid, M.; Singh, R.P.; Suman, R. Medical 4.0 technologies for healthcare: Features, capabilities, and applications. *Internet Things Cyber-Phys. Syst.* **2022**, *2*, 12–30. [CrossRef]
- DeTore, A.W. Medical informatics: An introduction to computer technology in medicine. Am. J. Med. 1988, 85, 399–403. [CrossRef] [PubMed]
- 5. Paik, S.-H.; Kim, D.-J. Smart Healthcare Systems and Precision Medicine. Adv. Exp. Med. Biol. 2019, 1192, 263–279. [CrossRef]
- Thirunavukarasu, R.; Gnanasambandan, R.; Gopikrishnan, M.; Palanisamy, V. Towards computational solutions for precision medicine based big data healthcare system using deep learning models: A review. *Comput. Biol. Med.* 2022, 149, 106020. [CrossRef]
- 7. Yu, G.; Chen, Z.; Wu, J.; Tan, Y. Medical decision support system for cancer treatment in precision medicine in developing countries. *Expert Syst. Appl.* **2021**, *186*, 115725. [CrossRef]
- 8. Altameem, A.; Kovtun, V.; Al-Ma'aitah, M.; Altameem, T.; Fouad, H.; Youssef, A.E. Patient's data privacy protection in medical healthcare transmission services using back propagation learning. *Comput. Electr. Eng.* **2022**, *102*, 108087. [CrossRef]
- Mosavi, N.S.; Santos, M.F. How Prescriptive Analytics Influences Decision Making in Precision Medicine. *Procedia Comput. Sci.* 2020, 177, 528–533. [CrossRef]
- 10. Kalra, D.; Ingram, D. Electronic Health Records. In *Information Technology Solutions for Healthcare. Health Informatics*; Zieliński, K., Duplaga, M., Ingram, D., Eds.; Springer: London, UK, 2006. [CrossRef]
- 11. Delen, D. Prescriptive Analytics The Final Frontier for Evidence-Based Management and Optimal Decision; Pearson Education Inc.: London, UK, 2020.
- 12. Tran, T.Q.B.; du Toit, C.; Padmanabhan, S. Artificial intelligence in healthcare—the road to precision medicine. *J. Hosp. Manag. Heal. Policy* **2021**, *5*, 29. [CrossRef]
- Kliem, P.S.; Tisljar, K.; Baumann, S.M.; Grzonka, P.; De Marchis, G.M.; Bassetti, S.; Bingisser, R.; Hunziker, S.; Marsch, S.; Sutter, R. First-Response ABCDE Management of Status Epilepticus: A Prospective High-Fidelity Simulation Study. J. Clin. Med. 2022, 11, 435. [CrossRef] [PubMed]
- 14. Mao, R.Q.; Lan, L.; Kay, J.; Lohre, R.; Ayeni, O.R.; Goel, D.P.; de Sa, D. Immersive Virtual Reality for Surgical Training: A Systematic Review. *J. Surg. Res.* 2021, 268, 40–58. [CrossRef]
- 15. Verma, D.; Bach, K.; Mork, P.J. Application of Machine Learning Methods on Patient Reported Outcome Measurements for Predicting Outcomes: A Literature Review. *Informatics* **2021**, *8*, 56. [CrossRef]
- 16. Yu, Y.; Li, M.; Liu, L.; Li, Y.; Wang, J. Clinical big data and deep learning: Applications, challenges, and future outlooks. *Big Data Min. Anal.* 2019, *2*, 288–305. [CrossRef]
- 17. Chakraborty, S.; Mallick, B.; Chakraborty, S. Mining of association rules for treatment of dental diseases. *J. Decis. Anal. Intell. Comput.* 2022, 2, 1–11. [CrossRef]
- Doğan, R.I.; Kim, S.; Chatr-Aryamontri, A.; Wei, C.-H.; Comeau, D.C.; Antunes, R.; Matos, S.; Chen, Q.; Elangovan, A.; Panyam, N.C.; et al. Overview of the BioCreative VI Precision Medicine Track: Mining protein interactions and mutations for precision medicine. *Database* 2019, 2019, bay147. [CrossRef]
- 19. Nguyen, D.; Long, T.; Jia, X.; Lu, W.; Gu, X.; Iqbal, Z.; Jiang, S. A feasibility study for predicting optimal radiation therapy dose distributions of prostate cancer patients from patient anatomy using deep learning. *Sci. Rep.* **2019**, *9*, 1076. [CrossRef]
- Du, Y.; Pan, Y.; Wang, C.; Ji, J. Biomedical semantic indexing by deep neural network with multi-task learning. *BMC Bioinform*. 2018, 19, 502. [CrossRef]
- 21. Izonin, I.; Tkachenko, R.; Duriagina, Z.; Shakhovska, N.; Kovtun, V.; Lotoshynska, N. Smart Web Service of Ti-Based Alloy's Quality Evaluation for Medical Implants Manufacturing. *Appl. Sci.* **2022**, *12*, 5238. [CrossRef]
- Tabares-Soto, R.; Orozco-Arias, S.; Romero-Cano, V.; Bucheli, V.S.; Rodríguez-Sotelo, J.L.; Jiménez-Varón, C.F. A comparative study of machine learning and deep learning algorithms to classify cancer types based on microarray gene expression data. *PeerJ Comput. Sci.* 2020, *6*, e270. [CrossRef]
- 23. Backenroth, D.; Chase, H.S.; Wei, Y.; Friedman, C. Monitoring prescribing patterns using regression and electronic health records. BMC Med. Inform. Decis. Mak. 2017, 17, 175. [CrossRef]
- 24. Arvanitis, A.; Furxhi, I.; Tasioulis, T.; Karatzas, K. Prediction of the effective reproduction number of COVID-19 in Greece. A machine learning approach using Google mobility data. *J. Decis. Anal. Intell. Comput.* **2021**, *1*, 1–21. [CrossRef]
- 25. Kasbekar, P.U.; Goel, P.; Jadhav, S.P. A Decision Tree Analysis of Diabetic Foot Amputation Risk in Indian Patients. *Front. Endocrinol.* **2017**, *8*, 25. [CrossRef]
- Tai, A.M.; Albuquerque, A.; Carmona, N.E.; Subramanieapillai, M.; Cha, D.S.; Sheko, M.; Lee, Y.; Mansur, R.; McIntyre, R.S. Machine learning and big data: Implications for disease modeling and therapeutic discovery in psychiatry. *Artif. Intell. Med.* 2019, 99, 101704. [CrossRef] [PubMed]

- Kesler, S.R.; Rao, A.; Blayney, D.W.; Oakley-Girvan, I.A.; Karuturi, M.; Palesh, O. Predicting Long-Term Cognitive Outcome Following Breast Cancer with Pre-Treatment Resting State fMRI and Random Forest Machine Learning. *Front. Hum. Neurosci.* 2017, 11, 555. [CrossRef] [PubMed]
- Chen, V.; Zhang, H. Depth importance in precision medicine (DIPM): A tree- and forest-based method for right-censored survival outcomes. *Biostatistics* 2020, 23, 157–172. [CrossRef]
- Rabcan, J.; Levashenko, V.; Zaitseva, E.; Kvassay, M. EEG Signal Classification Based On Fuzzy Classifiers. *IEEE Trans. Ind. Inform.* 2021, 18, 757–766. [CrossRef]
- 30. Zaitseva, E.; Levashenko, V.; Rabcan, J.; Krsak, E. Application of the Structure Function in the Evaluation of the Human Factor in Healthcare. *Symmetry* **2020**, *12*, 93. [CrossRef]
- 31. Abbod, M.F.; von Keyserlingk, D.G.; Linkens, D.A.; Mahfouf, M. Survey of utilisation of fuzzy technology in Medicine and Healthcare. *Fuzzy Sets Syst.* 2001, 120, 331–349. [CrossRef]
- Lin, I.; Loyola-González, O.; Monroy, R.; Medina-Pérez, M.A. A Review of Fuzzy and Pattern-Based Approaches for Class Imbalance Problems. *Appl. Sci.* 2021, 11, 6310. [CrossRef]
- Loh, H.W.; Ooi, C.P.; Seoni, S.; Barua, P.D.; Molinari, F.; Acharya, U.R. Application of explainable artificial intelligence for healthcare: A systematic review of the last decade (2011–2022). *Comput. Methods Programs Biomed.* 2022, 226, 107161. [CrossRef] [PubMed]
- 34. Prasanna, J.; Subathra, M.S.P.; Mohammed, M.A.; Damaševičius, R.; Sairamya, N.J.; George, S.T. Automated epileptic seizure detection in pediatric subjects of chb-mit eeg database—A survey. J. Pers. Med. 2021, 11, 1028. [CrossRef] [PubMed]
- Available online: https://uloz.to/file/W1WPexrq9k6Y/rats-data#!ZJWwMQR2ZGExBQqzLJH1AQLlBGuvZyx2nmyTAyxlEzu5 Lwt5Lj== (accessed on 1 February 2023).
- 36. Radha, P. Various Feature Selection Techniques in Type 2 Diabetic Patients for the Prediction of Cardiovascular Disease. *Int. J. Recent Innov. Trends Comput. Commun.* **2019**, *7*, 17–20. [CrossRef]
- Rabcan, J.; Zaitseva, E.; Levashenko, V.; Kvassay, M.; Surda, P.; Macekova, D. Fuzzy Decision Tree Based Method in Decision-Making of COVID-19 Patients' Treatment. *Mathematics* 2021, 9, 3282. [CrossRef]
- 38. Bodkhe, B.K.; Sood, S. Prediction of disease using fuzzy random forest. Int. J. Intell. Enterp. 2021, 8, 397–406. [CrossRef]
- Chatterjee, S.; Das, A. An ensemble algorithm using quantum evolutionary optimization of weighted type-II fuzzy system and staged Pegasos Quantum Support Vector Classifier with multi-criteria decision making system for diagnosis and grading of breast cancer. *Soft Comput.* 2023, 27, 7147–7178. [CrossRef]
- 40. Sharma, G.; Umapathy, K.; Krishnan, S. Trends in audio signal feature extraction methods. *Appl. Acoust.* **2020**, *158*, 107020. [CrossRef]
- 41. Polat, K.; Kara, S.; Güven, A.; Güneş, S. Usage of class dependency based feature selection and fuzzy weighted pre-processing methods on classification of macular disease. *Expert Syst. Appl.* **2009**, *36*, 2584–2591. [CrossRef]
- 42. Delgado, M.; Ruiz, M.D.; Sánchez, D.; Vila, M.A. Fuzzy quantification: A state of the art. *Fuzzy Sets Syst.* 2014, 242, 1–30. [CrossRef]
- 43. Yang, H.; Jiang, P.; Wang, Y.; Li, H. A fuzzy intelligent forecasting system based on combined fuzzification strategy and improved optimization algorithm for renewable energy power generation. *Appl. Energy* **2022**, 325, 119849. [CrossRef]
- 44. Glöckner, I. Fuzzy Quantifiers: A Computational Theory, Studies in Fuzziness and Soft Computing; Springer: New York, NY, USA, 2006; Volume 193.
- Nefti, S.; Oussalah, M. Probabilistic-fuzzy clustering algorithm. In Proceedings of the 2004 IEEE International Conference on Systems, Man and Cybernetics, Hague, The Netherlands, 10–13 October 2004; Volume 5, pp. 4786–4791.
- 46. Andreu-Perez, J.; Poon, C.C.Y.; Merrifield, R.D.; Wong, S.T.C.; Yang, G.-Z. Big Data for Health. *IEEE J. Biomed. Health Inform.* 2015, 19, 1193–1208. [CrossRef]
- 47. Laney, D. 3D data management: Controlling data volume, velocity, and variety. In *Application Delivery Strategies*; META Group Inc.: Stamford, UK, 2001.
- De Mauro, A.; Greco, M.; Grimaldi, M. A formal defnition of big data based on its essential features. *Libr. Rev.* 2016, 65, 122–135. [CrossRef]
- 49. Motai, Y.; Alam Siddique, N.; Yoshida, H. Heterogeneous data analysis: Online learning for medical-image-based diagnosis. *Pattern Recognit.* **2017**, *63*, *612–624*. [CrossRef]
- 50. Li, Z.; Qu, L.; Zhang, G.; Xie, N. Attribute selection for heterogeneous data based on information entropy. *Int. J. Gen. Syst.* 2021, 50, 548–566. [CrossRef]
- 51. Yue, L.; Tian, D.; Chen, W.; Han, X.; Yin, M. Deep learning for heterogeneous medical data analysis. *World Wide Web* 2020, 23, 2715–2737. [CrossRef]
- 52. Luo, Y.; Ahmad, F.S.; Shah, S.J. Tensor Factorization for Precision Medicine in Heart Failure with Preserved Ejection Fraction. J. *Cardiovasc. Transl. Res.* 2017, *10*, 305–312. [CrossRef] [PubMed]
- 53. Begoli, E.; Bhattacharya, T.; Kusnezov, D. The need for uncertainty quantification in machine-assisted medical decision making. *Nat. Mach. Intell.* **2019**, *1*, 20–23. [CrossRef]
- 54. Geiger, B.C.; Kubin, G. Information Loss in Deterministic Signal Processing Systems, 1st ed.; Springer International Publishing: New York, NY, USA, 2018.
- 55. Potapov, P. On the loss of information in PCA of spectrum-images. Ultramicroscopy 2017, 182, 191–194. [CrossRef]

- Yager, R.R. Toward a General Theory of Reasoning with Uncertainty, I: Nonspecificity and Fuzziness. *Int. J. Man-Mach. Stud.* 1986, 25, 613–631. [CrossRef]
- 57. Rokach, L. Using Fuzzy Logic in Data Mining. In *Data Mining and Knowledge Discovery Handbook*; Springer: New York, NY, USA, 2010; pp. 505–520.
- Burkov, A.; Paquet, S.; Michaud, G.; Valin, P. An Empirical Study of Uncertainty Measures in the Fuzzy Evidence Theory. In Proceedings of the 14th International Conference on Information Fusion, Chicago, IL, USA, 5–8 July 2011; pp. 453–460.
- 59. Zadeh, L.A. A computational approach to fuzzy quantifiers in natural languages. Comput. Math. Appl. 1983, 9, 149–184. [CrossRef]
- 60. Wolpert, D.; Macready, W. No free lunch theorems for optimization. *IEEE Trans. Evol. Comput.* **1997**, *1*, 67–82. [CrossRef]
- 61. Fernández-Delgado, M.; Cernadas, E.; Barro, S.; Amorim, D. Do we need hundreds of classifiers to solve real world classification problems? *J. Mach. Learn. Res.* 2014, *15*, 3133–3181.
- 62. Kaur, A.; Kaur, I. An empirical evaluation of classification algorithms for fault prediction in open source projects. J. King Saud Univ.—Comput. Inf. Sci. 2018, 30, 2–17. [CrossRef]
- 63. Anders, C.; Arnrich, B. Wearable electroencephalography and multi-modal mental state classification: A systematic literature review. *Comput. Biol. Med.* **2022**, *150*, 106088. [CrossRef]
- 64. Martinez-Ríos, E.A.; Bustamante-Bello, M.R.; Arce-Sáenz, L.A. A Review of Road Surface Anomaly Detection and Classification Systems Based on Vibration-Based Techniques. *Appl. Sci.* 2022, *12*, 9413. [CrossRef]
- 65. Houssein, E.H.; Hammad, A.; Ali, A.A. Human emotion recognition from EEG-based brain-computer interface using machine learning: A comprehensive review. *Neural Comput. Appl.* **2022**, *34*, 12527–12557. [CrossRef]
- 66. Arpitha, Y.; Madhumathi, G.L.; Balaji, N. Spectrogram analysis of ECG signal and classification efficiency using MFCC feature extraction technique. *J. Ambient. Intell. Humaniz. Comput.* **2021**, *13*, 757–767. [CrossRef]
- 67. Petrantonakis, P.C.; Hadjileontiadis, L.J. Emotion recognition from brain signals using hybrid adaptive filtering and higher order crossings analysis. *IEEE Trans. Affect. Comput.* 2010, *1*, 81–97. [CrossRef]
- 68. Peng, Y.; Qiu, T.; Wei, L. An approach to extracting graph kernel features from functional brain networks and its applications to the analysis of the noisy EEG signals. *Biomed. Signal Process. Control* **2023**, *80*, 104269. [CrossRef]
- 69. Alharbey, R.; Alsubhi, S.; Daqrouq, K.; Alkhateeb, A. The continuous wavelet transform using for natural ECG signal arrhythmias detection by statistical parameters. *Alex. Eng. J.* **2022**, *61*, 9243–9248. [CrossRef]
- Chen, J.; Bin, H.; Moore, P.; Zhang, X.; Ma, X. Electroencephalogram-based emotion assessment system using ontology and data mining techniques. *Appl. Soft Comput.* 2015, 30, 663–674. [CrossRef]
- 71. Chen, J.; Jiang, N.; Zhang, Y. A Common Spatial Pattern and Wavelet Packet Decomposition Combined Method for EEG-Based Emotion Recognition. *J. Adv. Comput. Intell. Inform.* **2019**, *23*, 274–281. [CrossRef]
- Jiang, C.; Li, Y.; Tang, Y.; Guan, C. Enhancing EEG-Based Classification of Depression Patients Using Spatial Information. *IEEE Trans. Neural Syst. Rehabil. Eng.* 2021, 29, 566–575. [CrossRef] [PubMed]
- 73. Subasi, A. Eeg signal classification using wavelet feature extraction and a mixture of expert model. *Expert Syst. Appl.* **2007**, *32*, 1084–1093. [CrossRef]
- Polat, K.; Güneş, S. Artificial immune recognition system with fuzzy resource allocation mechanism classifier, principal component analysis and FFT method based new hybrid automated identification system for classification of EEG signals. *Expert Syst. Appl.* 2008, 34, 2039–2048. [CrossRef]
- 75. Parhizkar, R.; Barbotin, Y.; Vetterli, M. Sequences with minimal time–frequency uncertainty. *Appl. Comput. Harmon. Anal.* 2015, 38, 452–468. [CrossRef]
- 76. Subha, D.P.; Joseph, P.K.; Acharya, R.; Lim, C.M. EEG Signal Analysis: A Survey. J. Med. Syst. 2010, 34, 195–212. [CrossRef]
- Li, M.; Chen, W.; Zhang, T. FuzzyEn-based features in FrFT-WPT domain for epileptic seizure detection. *Neural Comput. Appl.* 2019, *31*, 9335–9348. [CrossRef]
- Li, M.; Wang, R.; Yang, J.; Duan, L. An Improved Refined Composite Multivariate Multiscale Fuzzy Entropy Method for MI-EEG Feature Extraction. *Comput. Intell. Neurosci.* 2019, 2019, 7529572. [CrossRef]
- 79. Martinez, A.; Kak, A. PCA versus LDA. IEEE Trans. Pattern Anal. Mach. Intell. 2001, 23, 228–233. [CrossRef]
- Hesamian, G.; Akbari, M.G. Principal component analysis based on intuitionistic fuzzy random variables. *Comput. Appl. Math.* 2019, 38, 158. [CrossRef]
- 81. Subasi, A.; Gursoy, M.I. EEG signal classification using PCA, ICA, LDA and support vector machines. *Expert Syst. Appl.* **2010**, *37*, 8659–8666. [CrossRef]
- Akinola, O.O.; Ezugwu, A.E.; Agushaka, J.O.; Abu Zitar, R.; Abualigah, L. Multiclass feature selection with metaheuristic optimization algorithms: A review. *Neural Comput. Appl.* 2022, 34, 19751–19790. [CrossRef] [PubMed]
- Naheed, N.; Shaheen, M.; Khan, S.A.; Alawairdhi, M.; Khan, M.A. Importance of Features Selection, Attributes Selection, Challenges and Future Directions for Medical Imaging Data: A Review. Comput. Model. Eng. Sci. 2020, 125, 315–344. [CrossRef]
- Peng, H.; Long, F.; Ding, C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans. Pattern Anal. Mach. Intell.* 2005, 27, 1226–1238. [CrossRef]
- 85. Maldonado, J.; Riff, M.C.; Neveu, B. A review of recent approaches on wrapper feature selection for intrusion detection. *Expert Syst. Appl.* **2022**, *198*, 116822. [CrossRef]
- Yager, R.R. General multiple-objective decision functions and linguistically quantified statements. *Int. J. Man-Mach. Stud.* 1984, 21, 389–400. [CrossRef]

- 87. Ying, M. Linguistic quantifiers modeled by Sugeno integrals. Artif. Intell. 2006, 170, 581–606. [CrossRef]
- 88. Kupka, J. Some chaotic and mixing properties of fuzzified dynamical systems. Inf. Sci. 2014, 279, 642–653. [CrossRef]
- 89. Volna, E.; Jarusek, R.; Kotyrba, M.; Zacek, J. Training set fuzzification based on histogram to increase the performance of a neural network. *Appl. Math. Comput.* 2021, 398, 125994. [CrossRef]
- 90. Cánovas, J.; Kupka, J. Topological entropy of fuzzified dynamical systems. Fuzzy Sets Syst. 2011, 165, 37–49. [CrossRef]
- 91. Javadian, M.; Malekzadeh, A.; Heydari, G.; Shouraki, S.B. A clustering fuzzification algorithm based on ALM. *Fuzzy Sets Syst.* 2020, *389*, 93–113. [CrossRef]
- 92. Kaushal, M.; Lohani, Q.M.D. Generalized intuitionistic fuzzy c-means clustering algorithm using an adaptive intuitionistic fuzzification technique. *Granul. Comput.* 2022, 7, 183–195. [CrossRef]
- Bustamante, C.; Garrido, L.; Soto, R. Comparing Fuzzy Naive Bayes and Gaussian Naive Bayes for Decision Making in RoboCup 3D. In Proceedings of the Mexican International Conference on Artificial Intelligence, Apizaco, Mexico, 13–17 November 2006; pp. 237–247.
- 94. Kulkarni, A.D. Generating Classification Rules from Training Samples. Int. J. Adv. Comput. Sci. Appl. 2018, 9, 435. [CrossRef]
- De Carvalho, L.M.; Nassar, S.M.; De Azevedo, F.M.; De Carvalho, H.J.T.; Monteiro, L.L.; Rech, C.M.Z. A neuro-fuzzy system to support in the diagnostic of epileptic events and non-epileptic events using different fuzzy arithmetical operations. *Arq. de Neuro-Psiquiatr.* 2008, 66, 179–183. [CrossRef]
- 96. Yuan, Y.; Shaw, M.J. Induction of fuzzy decision trees. *Fuzzy Sets Syst.* **1995**, *69*, 125–139. [CrossRef]
- 97. Eusebi, P. Diagnostic Accuracy Measures. Cerebrovasc. Dis. 2013, 36, 267–272. [CrossRef]
- Levashenko, V.G.; Zaitseva, E.N. Usage of New Information Estimations for Induction of Fuzzy Decision Trees. In Intelligent Data Engineering and Automated Learning—IDEAL 2002; Yin, H., Allinson, N., Freeman, R., Keane, J., Hubbard, S., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2002; Volume 2412. [CrossRef]
- Andrzejak, R.G.; Lehnertz, K.; Mormann, F.; Rieke, C.; David, P.; Elger, C.E. Indications of nonlinear deterministic and finitedimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *Phys. Rev. E* 2001, 64 Pt 1, 061907. [CrossRef]
- 100. Takhar, A.; Surda, P.; Ahmad, I.; Amin, N.; Arora, A.; Camporota, L.; Denniston, P.; El-Boghdadly, K.; Kvassay, M.; Macekova, D.; et al. Timing of Tracheostomy for Prolonged Respiratory Wean in Critically Ill Coronavirus Disease 2019 Patients: A Machine Learning Approach. Crit. Care Explor. 2020, 2, e0279. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.