Article

Business Intelligence in Airline Passenger Satisfaction Study—A Fuzzy-Genetic Approach with Optimized Interpretability-Accuracy Trade-Off

Marian B. Gorzałczany *,[†] , **Filip Rudziński** [†] and **Jakub Piekoszewski** [†]

Department of Electrical and Computer Engineering, Kielce University of Technology, Al. 1000-lecia P.P. 7, 25-314 Kielce, Poland; f.rudzinski@tu.kielce.pl (F.R.); j.piekoszewski@tu.kielce.pl (J.P.)
* Correspondence: m.b.gorzalczany@tu.kielce.pl
† These authors contributed equally to this work.

Abstract: The main objective and contribution of this paper is the application of our knowledge-discovery business-intelligence technique (fuzzy rule-based classification systems) characterized by genetically optimized interpretability-accuracy trade-off (using multi-objective evolutionary optimization algorithms) to decision support related to airline passenger satisfaction problems. Recently published and accessible at Kaggle's repository airline passengers satisfaction data set containing 259,760 records is used in our experiments. A comparison of our approach with an alternative method (using SAS-system's accuracy-oriented prediction tools to determine the attribute importance hierarchy) is also performed showing the advantages of our method in terms of: (i) discovering the actual hierarchy of attribute significance for passenger satisfaction and (ii) knowledge-discovery system's interpretability-accuracy trade-off optimization. The main results and findings of our work include: (i) an introduction of the modern fuzzy-genetic business-intelligence solution characterized both by high interpretability and high accuracy to the airline passenger satisfaction decision support, (ii) an analysis of the effect of possible "overlapping" of some input attributes over the other ones in order to discover the real hierarchy of influence of particular input attributes upon the airline passengers satisfaction, and (iii) an extended cross-validation experiment confirming high effectiveness of our approach for different learning-test splits of the data set considered.

Keywords: business intelligence; airline passenger satisfaction; fuzzy rule-based systems; multi-objective evolutionary optimization; accuracy-interpretability trade-off optimization

1. Introduction

Business intelligence (BI), in general, aims at providing decision support—based on empirical information—for various business activities in different domains such as industry, science, technology, healthcare, commerce, defense, etc. [1,2]. BI is often used as an umbrella term that combines architectures, tools, databases, analytical tools, applications, and methodologies [3] summarizing a huge set of models and analytical methods such as reporting, data warehousing, and data mining [1]. Data about the considered business are the core elements of BI. Modern data mining (or knowledge discovery in data) tools rooted in the field of computational intelligence have given rise to knowledge-based decision support systems (equipped with knowledge bases and inference engines [4]) which can significantly enhance the formal apparatus of BI.

Knowledge-based BI approaches are well suited for decision support of business activities in aviation industry. It is due to the availability of huge amounts of data collected over time by airlines and airports on various aspects of their operation. A significant amount of domain knowledge is buried and hidden in such data. Effective knowledge-discovery approaches can reveal, in an automatic way, understandable and useful structures, trends and patterns in the considered data to improve accuracy and provide decision explanation in

aviation industry decision support. Many data mining techniques—including those rooted in the field of computational intelligence—have been applied to the airline data analysis—see, e.g., [5–10]. However, their essential drawback is their non-transparent, black-box, and accuracy-oriented nature, i.e., they do not provide any deeper (or any) explanations and justifications of the decisions made. This work is an attempt to address this problem by providing a solution characterized both by high interpretability and transparency as well as by high accuracy in the airline passenger satisfaction study.

The main goal and contribution of this paper is the application of our knowledge-discovery technique (fuzzy rule-based classification systems) characterized by genetically optimized interpretability-accuracy trade-off (see, e.g., [11–14] for details) to decision support related to airline passenger satisfaction problems. The quality of service and passenger satisfaction in the airline industry are increasingly recognized as critical factors of business performance and strategic tools for gaining competitive advantage [15]. In particular, we discover the hierarchy of influence of particular input attributes upon the airline passenger satisfaction. We also analyze the effect of possible "overlapping" of some input attributes over the other ones from the passenger satisfaction point of view. In our approach, measures of system's interpretability and accuracy are the separate performance indices and optimization objectives in designing fuzzy rule-based classifiers (FRBCs) from data. Due to the complementary/contradictory nature of both optimization objectives, multi-objective evolutionary optimization algorithms (MOEOAs) are employed in the process of the FRBC's structure and parameter optimization which is also equivalent to the FRBC's interpretability-accuracy trade-off optimization (see, e.g., [16] for review of related work and [17,18] for the single-objective optimization case).

First, the recently published Kaggle's airline passenger satisfaction data set containing as many as 259,760 records and used in our experiments is briefly characterized. Then, main components of our FRBCs and their MOEOA-based learning and optimization are outlined. For comparison purposes, two MOEOAs are used: our generalization (referred to as SPEA3 [19–21]) of the well-known Strength Pareto Evolutionary Algorithm 2 (SPEA2) [22] and SPEA2 itself. In turn, the afore outlined main goal of the paper, i.e., the application of our approach to the Kaggle's airline passenger satisfaction data and a comparative analysis with an alternative approach are presented and discussed.

2. Kaggle's Airline Passenger Satisfaction Data

As earlier mentioned, the recently published airline passenger satisfaction data set containing 259,760 records will be used in our experiments. It is a combination of two Excel data sets—referred to as "satisfaction.xlsx" and "satisfaction_2015.xlsx"—accessible at Kaggle's repository [23]; see also [24]. Each of them contains 129,880 records. Each record in both sets is characterized by 24 attributes. However, the attribute "Online support" occurring in the first set does not occur in the second set and the attribute "Inflight service" occurring in the second set does not occur in the first set. Following [24], both attributes were removed from the corresponding sets. Thus, each record of the final data set—obtained by merging the first and the second sets and referred to as the airline passenger satisfaction data set—is characterized by 23 attributes. Table 1 presents details of particular records of the final data set used in our experiments. The first attribute (i.e., "id" of a passenger—unique for each record) will not be used in our experiments. The second attribute, i.e., "satisfaction_v2" is the class attribute in our experiments. For better clarity, from now on its name will be replaced by the phrase "Passenger is" followed by the class label, i.e., either "neutral or dissatisfied" or "satisfied". The remaining 21 attributes are the input attributes. They include four numerical attributes, 13 qualitative ordinal attributes, and four qualitative nominal (sometimes referred to as categorical) attributes. Such data will be used in our experiments presented later in the paper. It is worth emphasizing that the data are almost perfectly balanced. As shown in Table 1, the classes "neutral or dissatisfied" and "satisfied" are represented by 49% and 51% of all data records, respectively.

Therefore, there is no bias when accuracy is used as one of performance measures of the obtained FRBCs.

Table 1. Details of particular records of the airline passenger satisfaction data set used in our experiments.

No.	Attribute Name	Attribute Type	Attribute Domain Details (·%—Percentage of the Overall Number of Samples)		
1.	id	numerical	Passenger's id (not used in our experiments)		
2.	satisfaction_v2 [*1]	class label	2 class labels: "neutral or dissatisfied" (49%) and "satisfied" (51%)		
3.	Gender	nominal [*2]	2 terms: "female" (51%) and "male" (49%)		
4.	Customer type	"	2 terms: "loyal customer" (82%) and "disloyal customer" (18%)		
5.	Age	numerical	integer numbers from 7 to 85 (average: 39.2, std. deviation: 15.1)		
6.	Type of travel	nominal [*2]	2 terms: "personal travel" (31%) and "business travel" (69%)		
7.	Class	"	3 terms: "eco" (46%), "business" (48%), and "eco plus" (7%)		
8.	Flight distance	numerical	integer numbers from 31 to 6907 (average: 1590.3, std. deviation: 1082.5)		
9.	Inflight WiFi service	ordinal [*3]	'0' (1.6%)	'1' (14.3%)	'2' (22.9%)
			'3' (23.0%)	'4' (21.7%)	'5' (16.6%)
10.	Departure/Arrival time convenient	"	'0' (5.1%)	'1' (15.5%)	'2' (17.1%)
			'3' (17.5%)	'4' (23.7%)	'5' (21.1%)
11.	Ease of online booking	"	'0' (2.2%)	'1' (13.6%)	'2' (19.2%)
			'3' (20.3%)	'4' (24.8%)	'5' (19.8%)
12.	Gate location	"	'0' (0%)	'1' (17.2%)	'2' (18.8%)
			'3' (26.7%)	'4' (23.3%)	'5' (14.1%)
13.	Food and drink	"	'0' (2.3%)	'1' (14.3%)	'2' (21%)
			'3' (21.5%)	'4' (22.2%)	'5' (18.6%)
14.	Online boarding	"	'0' (1.2%)	'1' (11%)	'2' (15.6%)
			'3' (22.3%)	'4' (28.4%)	'5' (21.6%)
15.	Seat comfort	"	'0' (1.8%)	'1' (13.9%)	'2' (18.2%)
			'3' (20.2%)	'4' (26.2%)	'5' (19.6%)
16.	Inflight entertainment	"	'0' (1.2%)	'1' (10.6%)	'2' (15.8%)
			'3' (18.5%)	'4' (30.3%)	'5' (23.6%)
17.	On-board service	"	'0' (0%)	'1' (10.8%)	'2' (13.7%)
			'3' (21.4%)	'4' (30.6%)	'5' (23.6%)
18.	Leg room service	"	'0' (0.4%)	'1' (9.3%)	'2' (17.8%)
			'3' (18.3%)	'4' (29.1%)	'5' (25.1%)
19.	Baggage handling	"	'0' (0%)	'1' (6.5%)	'2' (10.7%)
			'3' (19.4%)	'4' (36.6%)	'5' (26.8%)
20.	Checkin service	"	'0' (0%)	'1' (12.1%)	'2' (12.2%)
			'3' (27.3%)	'4' (28%)	'5' (20.4%)
21.	Cleanliness	"	'0' (0%)	'1' (9.4%)	'2' (12.9%)
			'3' (21%)	'4' (31.9%)	'5' (24.8%)
22.	Departure delay in minutes	numerical	integer numbers from 0 to 1592 (average: 14.6, std. deviation: 39.0)		
23.	Arrival delay in minutes	"	integer numbers from 0 to 1584 (average: 14.9, std. deviation: 39.4)		

[*1] henceforward, for better clarity the name "satisfaction_v2" will be replaced by "Passenger is" (see comments in the paper), [*2] sometimes referred to as categorical, [*3] satisfaction level of the passenger: "0" (no answer), "1" (the minimal one), ..., "5" (the maximal one).

3. Methodology: An Outline of Main Components of the Proposed FRBCs and Their MOEOA-Based Learning and Optimization

For convenience of the reader, in this section we briefly outline main components of our approach—see [11–13,19] for details and discussion. In general, we consider FRBC with n input attributes $x_1, x_2, \ldots, x_n$ (including both numerical and qualitative ones) and an output—a fuzzy set over the set $Y = \{y_1, y_2, \ldots, y_c\}$ of c class labels.

Learning data set **L**: it is a basis for the FRBC's design from data and contains K input-output samples:

$$L = \{x_k^{(lrn)}, y_k^{(lrn)}\}_{k=1}^{K}, \tag{1}$$

where $x_k^{(lrn)} = (x_{1k}^{(lrn)}, x_{2k}^{(lrn)}, \ldots, x_{nk}^{(lrn)}) \in X = X_1 \times X_2 \times \cdots \times X_n$ ($\times$ stands for Cartesian product of ordinary sets) is the set of input attributes and $y_k^{(lrn)}$ is the corresponding class label ($y_k^{(lrn)} \in Y$) for the k-th data sample, $k = 1, 2, \ldots, K$.

Attribute representation: each numerical attribute x_i, $i \in \{1, 2, \ldots, n\}$ is represented by a_i fuzzy sets $A_{ik_i} \in F(X_i)$ ($F(X_i)$ denotes a family of all fuzzy sets defined in the universe X_i), $k_i = 1, 2, \ldots, a_i$. A_{i_1} is an S-type fuzzy set (representing linguistic term "*Small*"), A_{ia_i} is an L-type set (representing linguistic term "*Large*"), and $A_{i_2}, A_{i_3}, \ldots, A_{i,a_i-1}$ are M-type sets (representing linguistic terms "*Medium 1*", "*Medium 2*", ... , "*Medium $a_i - 2$*"). For simplicity, A_{ik_i}s denote also the corresponding linguistic terms. Trapezoidal membership functions of S-, M-, and L-type fuzzy sets are shown in Figure 1. Each qualitative attribute x_i, $i \in \{1, 2, \ldots, n\}$ ($x_i \in X_i = \{x_{i1}, x_{i2}, \ldots, x_{ia_i}\}$) is characterized by a_i fuzzy singletons $A_{ik_i} = A_{ik_i}^{(singl.)}$, $k_i = 1, 2, \ldots, a_i$ defined for particular "values" x_{ik_i} of x_i as follows: $\mu_{A_{ik_i}^{(singl.)}}(x_i) = 1$ for $x_i = x_{ik_i}$ and 0 elsewhere. Similarly, class labels $y_{j^{(r)}}$ in particular rules (2) ($j^{(r)} \in \{1, 2, \ldots, c\}$) are represented by appropriate fuzzy singletons $B_{j^{(r)}}^{(singl.)}$ with the following membership functions: $\mu_{B_{j^{(r)}}^{(singl.)}}(y) = 1$ for $y = y_{j^{(r)}}$ and 0 elsewhere.
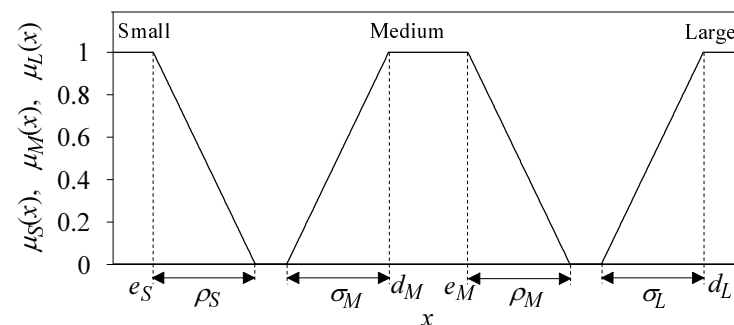
Figure 1. Trapezoidal membership functions of S-type, M-type, and L-type fuzzy sets and their parameters.

FRBC's knowledge base: it contains R genetically optimized fuzzy rules discovered in the learning data set (1). We introduced the following form of the r-th rule, $r = 1, 2, \ldots, R$ (R changes during the learning process):

$$\begin{aligned}
&\textbf{IF } \left[x_1 \text{ is } [\text{not}]_{(sw_1^{(r)}<0)} \ A_{1,|sw_1^{(r)}|}\right]_{(sw_1^{(r)} \neq 0)} \textbf{ AND...AND} \\
&\left[x_n \text{ is } [\text{not}]_{(sw_n^{(r)}<0)} \ A_{n,|sw_n^{(r)}|}\right]_{(sw_n^{(r)} \neq 0)} \\
&\textbf{THEN } y \text{ is } B_{j^{(r)}}^{(singl.)},
\end{aligned} \tag{2}$$

where: (i) $[expression]_{(condition)}$ in (2) denotes conditional inclusion of $[expression]$ into a given rule if and only if $(condition)$ is fulfilled, (ii) $|\cdot|$ returns the absolute value,

and (iii) $sw_i^{(r)}$ is a switch which controls the presence/absence of the i-th input attribute in the r-th rule, $i = 1, 2, \ldots, n$.

$sw_i^{(r)} \in \{0, \pm 1, \pm 2, \ldots, \pm a_i\}$, where a_i is the number of fuzzy sets (linguistic terms) defined for the i-th attribute. For $sw_i^{(r)} = 0$, the i-th attribute is excluded from (not active in) that rule, whereas for $sw_i^{(r)} > 0$ the component $[x_i$ is $A_{ik_i}]$ ($k_i = |sw_i^{(r)}|$) is included (active) and for $sw_i^{(r)} < 0$ the component $[x_i$ is not $A_{ik_i}]$ is used in that rule (not $A_{ik_i} = \bar{A}_{ik_i}$ and $\mu_{\bar{A}_{ik_i}}(x_i) = 1 - \mu_{A_{ik_i}}(x_i)$; $\mu_{A_{ik_i}}(x_i)$ and $\mu_{\bar{A}_{ik_i}}(x_i)$ are membership functions of fuzzy sets A_{ik_i} and $\bar{A}_{ik_i}$).

In our approach two entities, i.e., rule base structure RB and data base DB represent FRBC's knowledge base. We introduce direct and computationally efficient RB's representation in the following form:

$$RB = \{sw_1^{(r)}, sw_2^{(r)}, \ldots, sw_n^{(r)}, j^{(r)}\}_{r=1}^{R}. \tag{3}$$

In turn, DB contains: (i) a tunable part, i.e., parameters of fuzzy sets representing numerical attributes and (ii) a non-tunable part, i.e., domains of qualitative attributes $X_i = \{x_{i1}, x_{i2}, \ldots, x_{ia_i}\}$, $i \in \{1, 2, \ldots, n\}$ and the set of class labels $Y = \{y_1, y_2, \ldots, y_c\}$.

We developed original crossover and mutation operators for the transformation of the RB population as well as we adopted some specialized crossover and mutation operators for the DBs processing—see [11–13] for details.

Evolutionary optimization objectives—FRBC's accuracy: the accuracy measure (subject to maximization) is defined as follows:

$$Q_{ACC}^{(lrn)} = 1 - Q_{RMSE}^{(lrn)}, \tag{4}$$

where

$$Q_{RMSE}^{(lrn)} = \sqrt{\frac{1}{Kc} \sum_{k=1}^{K} \sum_{j=1}^{c} \left[\mu_{B_k^{(singl.)(lrn)}}(y_j) - \mu_{B_k'}(y_j) \right]^2}. \tag{5}$$

$Q_{RMSE}^{(lrn)} \in [0, 1]$; B_k' (characterized by its membership function $\mu_{B_k'}(y)$) is the fuzzy-set response of system (2) for the learning data sample $x_k^{(lrn)}$, and $B_k^{(singl.)(lrn)}$ (represented by its membership function $\mu_{B_k^{(singl.)(lrn)}}(y) = 1$ for $y = y_k^{(lrn)}$ and 0 elsewhere) is the desired fuzzy-singleton response for that sample.

Evolutionary optimization objectives—FRBC's interpretability: we use the notion interpretability in a broader sense including not only semantic aspects of FRBC but also its complexity. We proposed the following measure (subject to maximization) for the evaluation of the FRBC's complexity-related interpretability:

$$Q_{INT} = 1 - Q_{CPLX}, \tag{6}$$

where

$$Q_{CPLX} = \frac{Q_{RINP} + Q_{INP} + Q_{FS}}{3}, \tag{7}$$

and

$$Q_{RINP} = \frac{1}{R} \sum_{r=1}^{R} \frac{n_{INP}^{(r)} - 1}{n - 1}, \; n > 1, \tag{8}$$

$$Q_{INP} = \frac{n_{INP} - 1}{n - 1}, \; Q_{FS} = \frac{n_{FS} - 1}{\sum_{i=1}^{n} a_i - 1}, \; n > 1. \tag{9}$$

The FRBC's complexity measure Q_{CPLX} (7) ($Q_{CPLX} \in [0, 1]$; 0 and 1 represent minimal and maximal complexities, respectively) is an average of three sub-indices that measure an average complexity of particular rules Q_{RINP} (8) and the complexity of the whole system in

terms of its active inputs Q_{INP} (9) and active fuzzy sets Q_{FS} (9). $n_{INP}^{(r)}$ in (8) is the number of active input attributes in the *r*-th rule. n_{INP} and n_{FS} in (9) are the numbers of active inputs and fuzzy sets (linguistic terms), respectively, in the whole system.

　　In turn, the FRBC's semantic-related interpretability is addressed by us by implementing strong fuzzy partitions (SFPs) [25] of domains of all numerical attributes. SFPs are fuzzy partitions in which the sum of the values of all membership functions for any domain value is equal to 1. SFPs satisfy the desired semantics-related interpretability demands [25]. Simple and computationally efficient implementation of SFP requirements for trapezoidal membership functions can be formulated as follows (see Figure 2 for three-set SFP of x_i):

$$\sigma_{ik_i} = \rho_{i,k_i-1} = d_{ik_i} - e_{i,k_i-1}, \quad k_i = 2, 3, \ldots, a_i \tag{10}$$

and, obviously,

$$e_{i1} \le d_{i2} \le e_{i2} \le \cdots \le d_{i,a_i-1} \le e_{i,a_i-1} \le d_{ia_i},$$
$$i = 1, 2, \ldots, n. \tag{11}$$
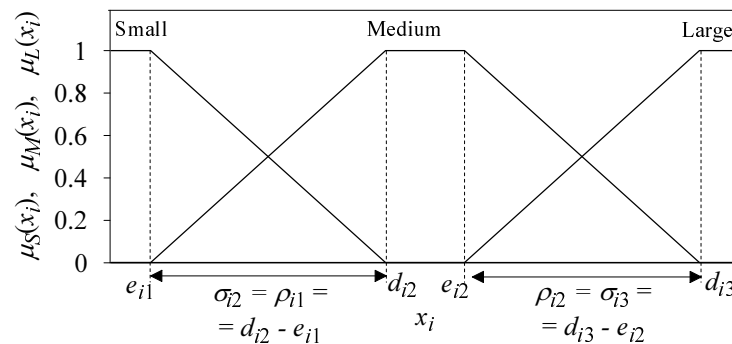
Figure 2. Implementation of three-fuzzy-set SFP.

　　MOEOAs used: as already mentioned in the Introduction of the paper, for comparison purposes, two MOEOAs are used in our experiments. They include the well-known SPEA2 method and our generalization of SPEA2 referred to as SPEA3. Three indices are usually applied to evaluate and compare different MOEOAs [25]. These indices include: (i) the accuracy of non-dominated solutions obtained, (ii) the spread of solutions in the solution set, and (iii) the distribution of solutions in the solution set. The accuracy represents the closeness of the generated non-dominated solutions to Pareto-optimal solutions (if they are available) or to reference solutions. The spread of solutions in the solution set—measured by the distance between extreme solutions in the set—represents how well the generated solutions arrive at the extrema of the Pareto-optimal or reference solution set. In turn, the distribution of solutions in the solution set represents how evenly the solutions are distributed along the approximation of Pareto-optimal/reference front in the objective space. Therefore, a set of solutions characterized by a higher accuracy, a higher spread, and a better-balanced distribution outperforms the alternative sets of solutions.

　　An appropriate MOEOA for our experiments should have the ability to generate solution set of possibly high accuracy, spread and distribution balance. It is directly related to the analysis of possible "overlapping" of some of input attributes over the other ones in the airline passenger satisfaction data (see the next section for a detailed presentation). Among the traditional MOEOAs, SPEA2 and NSGA-II (Nondominated Sorting Genetic Algorithm II) [26] are the best known. Since SPEA2 has a higher ability to generate solution sets of better-balanced distribution ([22,26]), we selected SPEA2 for further improvement of the spread and distribution of generated solutions.

　　The essence of the proposed SPEA2's generalization (referred to as SPEA3) consists in replacing its environmental selection procedure by our original algorithm, which improves the spread, distribution balance and diversity of generated solutions. The environmental

selection, in general, creates a collection of the best solutions out of all solutions obtained so far and keeps them in an external archive. Table 2 presents differences between environmental selection procedure of SPEA2 and our original environmental selection procedure implemented in SPEA3.

Table 2. Differences between environmental selection procedures of SPEA2 and SPEA3.

Environmental Selection Procedure of SPEA2:	Our Original Environmental Selection Procedure Implemented in SPEA3:
(a) immediately (i.e., in each generation of the optimization process) copies to the archive all available non-dominated solutions (if their number is lesser than the archive size, the best dominated solutions from the current population are also selected and copied to the archive to fully fill it),	(a) gradually (i.e., in the subsequent generations of the optimization process) fills the archive with only such non-dominated solutions which ensure the best balance of distances between neighboring solutions in the archive,
(b) truncates overfilled archive by removing from it redundant solutions characterized by the shortest distances to other solutions.	(b) truncates archive by removing from it solutions which have been dominated by any new solutions that have appeared in the current population,
	(c) gradually exchanges some non-dominated solutions between the archive and the current population in order to maximize the sum of distances between the nearest neighboring solutions in the archive.

Concluding:

(**a**)   in our original approach implemented in SPEA3, the complementary operations of increasing and reducing the archive lead to obtaining the best available distribution balance and spread of solutions belonging to Pareto-front approximation (see [19] for a detailed presentation and [20] for a discussion), whereas

(**b**)   in SPEA2 only the truncation procedure (if activated) contributes to improving the distribution balance and diversity of the final set of solutions (not addressing, however, the problem of improving the spread of solutions) and giving, in general, worse results than its counterpart of SPEA3.

4. Experiments (Application to Kaggle's Airline Passenger Satisfaction Data) and Discussion

First, we would like to reveal some details of the operation of our approach to design FRBCs from the considered Kaggle's airline data. For this purpose, the MOEOA-based genetic learning and optimization experiments for a single learning-test data split are presented and discussed. The ratio of split of the original data was 1:9, i.e., only 10% of the whole data set (preserving the class proportions) was used as the learning data to build the system whereas the remaining 90% of the original data were used for testing purposes. The learning-test data split applied in our experiments is widely used in processing data sets with large number of data samples (in our case, we dealt with as many as 259,760 data samples). Such a data split not only reduces the computational complexity of the learning procedure but, more importantly, formulates much higher demands for the classification technique since "generally, the larger the training dataset, the better the classification performance regardless of which classification algorithm is used" (quotation from [27]). Therefore, the assumed data split poses a significant challenge for the system's design technique. For comparison purposes, Figure 3 presents two 10-element collections of non-dominated solutions (i.e., optimized FRBCs) obtained in a final generation of a single run of our FRBC's design technique using, independently, our SPEA3 and SPEA2 methods. Both collections represent the best available approximations of Pareto-optimal solutions generated by our SPEA3 and SPEA2, respectively. Particular solutions from a given front were characterized by different levels of optimized accuracy–interpretability trade-off

allowing the user to select a single solution (a specific FRBC) characterized by a desired level of compromise between the accuracy and interpretability. Figure 3 shows that our SPEA3-based approach outperforms the SPEA2-based one by generating the collection of solutions characterized by much-better-balanced distribution in the objective space (the solutions are distributed along the front in a much more even way). The accuracy- and interpretability-related numerical details of all SPEA3-based solutions from Figure 3 are collected in Table 3, in which $n_{INP/R}$ is the number of input attributes per rule, $ACC^{(lrn)}$ and $ACC^{(tst)}$ are the percentages of correct decisions in the learning and test sets, respectively (the remaining parameters were defined earlier in the paper). Fuzzy rule bases of the first seven solutions from Figure 3 and Table 3 are presented in Tables 4 and 5 (membership functions of fuzzy sets used in those rule bases are shown in Figure 4). The remaining three solutions Nos. 8, 9, and 10 from Figure 3 and Table 3 have not been included in Table 5 because—comparing with solution No. 7—they provide very small increases in the test-data-accuracy $ACC^{(tst)}$ (by 0.1%, 0.4%, and 0.5%, respectively—see Table 3) despite significant increase in their complexity—see their interpretability measures in Table 3.
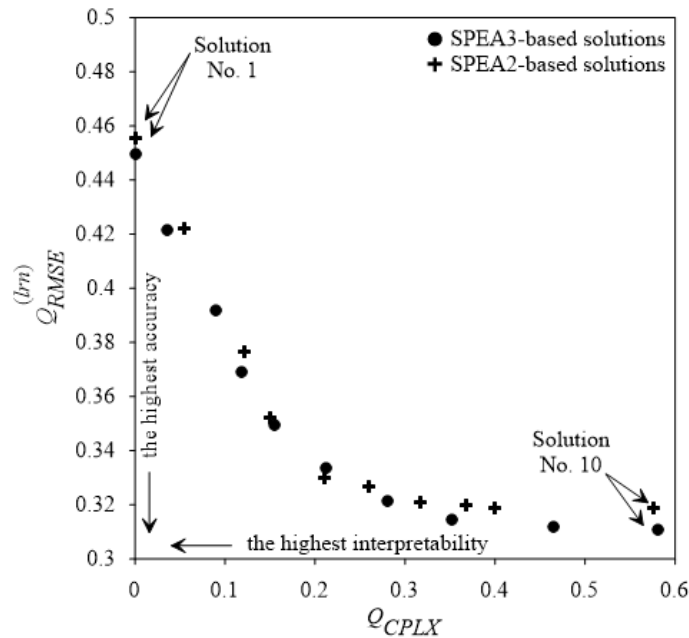
Figure 3. The best Pareto-front approximations generated by our SPEA3 and SPEA2 for the considered Kaggle's airline data.

Table 3. Interpretability and accuracy measures of SPEA3-based solutions from Figure 3.

| No. | Objective Function Complements | | Interpretability Measures | | | | Accuracy Measures | |
	$1 - Q_{INT} = Q_{CPLX}$	$1 - Q_{ACC}^{(lrn)} = Q_{RMSE}^{(lrn)}$	R	n_{INP}	n_{FS}	$n_{INP/R}$	$ACC^{(lrn)}$	$ACC^{(tst)}$
1.	0	0.4498	3	1	3	1	75.1%	75.2%
2.	0.035	0.4214	5	2	6	1.4	77.2%	77.2%
3.	0.0892	0.3917	7	4	11	1.6	80.5%	80.5%
4.	0.1187	0.3693	8	5	13	1.7	82.8%	82.8%
5.	0.1555	0.3497	9	6	15	2.2	84.6%	84.5%
6.	0.2113	0.3336	11	8	18	2.4	86.6%	86.4%
7.	0.2808	0.3216	17	11	25	2.2	88.3%	88.1%
8.	0.3514	0.3143	17	13	31	3.1	88.5%	88.2%
9.	0.4654	0.3119	21	17	37	3.6	88.8%	88.5%
10.	0.5805	0.311	27	21	45	4.2	88.9%	88.6%

Table 4. Fuzzy rule bases for SPEA3-based solutions (FRBCs) Nos. 1–5 from Figure 3 and Table 3.

No.	Fuzzy Classification Rules

Solution No. 1 ($ACC^{(lrn)} = 75.1\%$, $ACC^{(tst)} = 75.2\%$):

1. **IF** Inflight entertainment is no_answer or low **THEN** Passenger is neutral or dissatisfied
2. **IF** Inflight entertainment is medium **THEN** Passenger is neutral or dissatisfied
3. **IF** Inflight entertainment is high **THEN** Passenger is satisfied

Solution No. 2 ($ACC^{(lrn)} = 77.2\%$, $ACC^{(tst)} = 77.2\%$):

1. *This rule is an extension of rule No. 1 from Solution No. 1:*
   IF Inflight entertainment is no_answer or low **AND** Seat comfort is low or medium **THEN** Passenger is neutral or dissatisfied
2. *This rule is an extension of rule No. 2 from Solution No. 1:*
   IF Inflight entertainment is medium **AND** Seat comfort is low or medium **THEN** Passenger is neutral or dissatisfied
3. *This rule is the same as rule No. 3 from Solution No. 1.*
4. **IF** Seat comfort is no_answer **THEN** Passenger is satisfied
5. **IF** Seat comfort is high **THEN** Passenger is satisfied

Solution No. 3 ($ACC^{(lrn)} = 80.5\%$, $ACC^{(tst)} = 80.5\%$):

1. *This rule is the same as rule No. 1 from Solution No. 1.*
2. *This rule is the same as rule No. 2 from Solution No. 2.*
3. *This rule is an extension of rule No. 3 from Solution No. 2:*
   IF Inflight entertainment is high **AND** Type of travel is business travel **THEN** Passenger is satisfied
4. *This rule is the same as rule No. 4 from Solution No. 2.*
5. *This rule is an extension of rule No. 5 from Solution No. 2:*
   IF Seat comfort is high **AND** Inflight WiFi service is high **THEN** Passenger is satisfied
6. **IF** Inflight WiFi service is low or medium **AND** Type of travel is personal travel **THEN** Passenger is neutral or dissatisfied
7. **IF** Inflight WiFi service is no_answer **THEN** Passenger is satisfied

Solution No. 4 ($ACC^{(lrn)} = 82.8\%$, $ACC^{(tst)} = 82.8\%$):

1–2. *These rules are the same as rules Nos. 1 and 2 from Solution No. 2.*
3. *This rule is an extension of rule No. 3 from Solution No. 3:*
   IF Inflight entertainment is high **AND** Type of travel is business travel **AND** Customer type is loyal customer **THEN** Passenger is satisfied
4–7. *These rules are the same as rules Nos. 4–7 from Solution No. 3.*
8. **IF** Customer type is disloyal customer **AND** Inflight WiFi service is low (or medium) **THEN** Passenger is neutral or dissatisfied

Solution No. 5 ($ACC^{(lrn)} = 84.6\%$, $ACC^{(tst)} = 84.5\%$):

1–4. *These rules are the same as rules Nos. 1–4 from Solution No. 4.*
5. *This rule is the second extension of rule No. 5 from Solution No. 2:*
   IF Seat comfort is high **AND** Flight distance is long **THEN** Passenger is satisfied
6. *This rule is an extension of rule No. 6 from Solution No. 4:*
   IF Inflight WiFi service is low or medium **AND** Type of travel is personal travel **AND** Flight distance is short **THEN** Passenger is neutral or dissatisfied
7. *This rule is the same as rule No. 7 from Solution No. 4.*
8. **IF** Inflight WiFi service is high **AND** Flight distance is short **THEN** Passenger is satisfied
9. **IF** Inflight entertainment is high **AND** Inflight WiFi service is low or medium **AND** Customer type is disloyal customer **AND** Flight distance is short **THEN** Passenger is neutral or dissatisfied

Table 5. (Continuation of Table 4) Fuzzy rule bases for SPEA3-based solutions (FRBCs) Nos. 6 and 7 from Figure 3 and Table 3.

No.	Fuzzy Classification Rules

Solution No. 6 ($ACC^{(lrn)} = 86.6\%$, $ACC^{(tst)} = 86.4\%$):

1–5. *These rules are the same as rules Nos. 1–5 from Solution No. 5.*

6. *This rule is an extension of rule No. 6 from Solution No. 5:*
   IF Inflight WiFi service is low or medium **AND** Type of travel is personal travel **AND** Flight distance is short **AND** Customer type is loyal customer **THEN** Passenger is neutral or dissatisfied

7–8. *These rules are the same as rules Nos. 7–8 from Solution No. 5.*

9. **IF** Inflight WiFi service is low or medium **AND** Flight distance is short **AND** Customer type is disloyal customer **THEN** Passenger is neutral or dissatisfied

10. **IF** Inflight WiFi service is low or medium **AND** Customer type is loyal customer **AND** Ease of online booking is high **AND** Leg room service is high **THEN** Passenger is satisfied

11. **IF** Inflight WiFi service is low or medium **AND** Type of travel is personal travel **AND** Leg room service is low or medium **THEN** Passenger is neutral or dissatisfied

Solution No. 7 ($ACC^{(lrn)} = 88.3\%$, $ACC^{(tst)} = 88.1\%$):

1–5. *These rules are the same as rules Nos. 1–5 from Solution No. 6.*

6. *This rule is an extension of rule No. 6 from Solution No. 4:*
   IF Inflight WiFi service is low or medium **AND** Type of travel is personal travel **AND** Ease of online booking is no_answer or low **THEN** Passenger is neutral or dissatisfied

7–9. *These rules are the same as rules Nos. 7–9 from Solution No. 6.*

10. *This rule is an extension of rule No. 1 from Solution No. 1:*
    IF Inflight entertainment is no_answer or low **AND** Seat comfort is no_answer **THEN** Passenger is satisfied

11. *This rule is an extension of rule No. 6 from Solution No. 4:*
    IF Inflight WiFi service is low or medium **AND** Type of travel is personal travel **AND** Ease of online booking is medium **THEN** Passenger is neutral or dissatisfied

12. **IF** Inflight WiFi service is no_answer **AND** Ease of online booking is no_answer or low **THEN** Passenger is satisfied

13. **IF** On-board service is no_answer or low **AND** Seat comfort is no_answer **THEN** Passenger is satisfied

14. **IF** On-board service is no_answer or low **AND** Seat comfort is low or medium **THEN** Passenger is neutral or dissatisfied

15. **IF** On-board service is medium **AND** Leg room service is no_answer **THEN** Passenger is satisfied

16. **IF** On-board service is high **AND** Ease of online booking is high **AND** Leg room service is high **THEN** Passenger is satisfied

17. **IF** Seat comfort is low or medium **AND** Baggage handling is low or medium **AND** Checkin service is no_answer or low **THEN** Passenger is neutral or dissatisfied
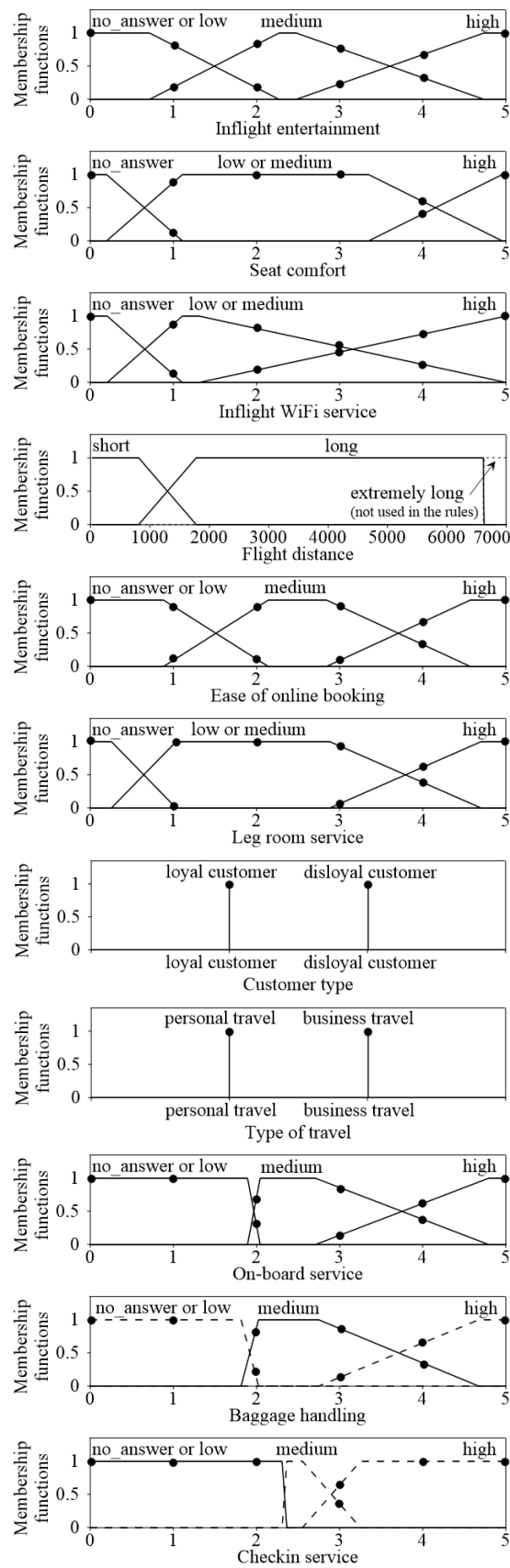
Figure 4. Final shapes of membership functions of input attributes occurring in fuzzy rule bases of Tables 4 and 5.

Tables 4 and 5 also reveal an interesting regularity, i.e., the fuzzy rule base of the solution No. *i* contains some rule(s) or extension(s) of some rule(s) from the solution No. $i-1$ ($i = 2, 3, \dots$). Therefore, if the higher accuracy is required then our approach adds some additional fuzzy rules or extends the already existing rules to provide a more detailed (and thus, more accurate) description of the considered classification problem. Such regularity also confirms an internal integrity of our approach. The considered regularity is illustrated—from a bit different angle—in Tables 6–8 (see, first, Part A of Table 6), in which each black square denotes the presence of a given input attribute in the fuzzy rule base of a given solution (FRBC). $ACC_1^{(tst)}$ is the test-data-accuracy of the system exclusively based on the most significant input attribute (attribute "Inflight entertainment", system (solution) No. 1, accuracy 75.2%). $\Delta ACC_j^{(tst)}$, $j = 2, 3, \dots$ is the accuracy increase following the inclusion of 2nd, 3rd,…most significant attribute into the system. For instance, the inclusion of "Seat comfort" (2nd most significant attribute) yields +2.0% increase in test-data accuracy and is related to system (solution) No. 2. In turn, the inclusion of "Type of travel" and "Inflight WiFi service" (3rd most significant attributes) gives further 3.2% test-data-accuracy-increase and is related to solution No. 3, etc.

The above presented reasoning is correct provided that there is no "overlapping" of some of input attributes over the other ones in the airline passenger satisfaction data. In order to verify that aspect of airline data, we remove from the original data set the so-far most significant attribute (i.e., "Inflight entertainment") and we repeat, in an analogous way, the learning experiment. Its results are presented in Part B of Table 6 giving "Online boarding" attribute the most significant place. Clearly "Online boarding" which occupied a low position in experiment of Part A was "overlapped" by "Inflight-entertainment". In the next step, we remove "Online boarding" from the present data set and repeat the learning experiment—see Part C of Table 6—obtaining "Ease of online booking" as the most significant attribute at this stage. It occupied second position in experiments of Part B. Therefore, we can conclude that it was not "overlapped" by "Online boarding". We repeat analogous experiments several times, i.e., removing the most significant attribute at a given stage and repeating the learning process on the reduced data—see Parts D, E, F of Table 7, and G, H, I of Table 8. In such a way, we arrive to the final hierarchy of input attribute significance from the perspective of the airline passenger satisfaction. It is shown in the left part of Table 9. $ACC_1^{(tst)}$ of Table 9 means the same as in Tables 6–8, i.e., the test-data-accuracy of the system exclusively based on a single attribute (listed to the left of $ACC_1^{(tst)}$ in Table 9).

The right part of Table 9 presents the results of alternative approach by Patlolla [24]. The paper [24], to our knowledge, is the only available now reference addressing the considered data set (most probably due to the fact that the data set has been published most recently). The approach of [24] uses SAS system to calculate the mean values of particular input attributes separately for both classes of passengers (satisfied and neutral or dissatisfied) and to build decision tree to determine the attribute importance hierarchy. Figure 5 summarizes, in a graphical form, details of our approach (the left part of a block scheme of Figure 5) and the alternative Patlolla's method [24] (the right part of that block scheme) for the purpose of their comparative analysis regarding the determination of the final hierarchy of input attribute significance. Data preprocessing, data partitioning, learning process, selection of the most significant input attributes, and determining the final hierarchy of attribute significance are the main stages of performing that task. Since 9 most significant input attributes were selected in the alternative work [24], we also select the same number of the most significant attributes. Although both approaches select "Inflight entertainment" as the most significant attribute (as shown in Table 9), our approach provides much deeper insight into the mechanisms encoded in the considered airline data set (see collections of easily-interpretable linguistic, fuzzy classification rules presented in Tables 4 and 5).

An important part of our work is the cross-validation experiment with 1:9 learning-test data split ratio. Each single learning experiment starts with generation of a Pareto-front

approximation. Then, a single solution characterized, first, by the highest test-data accuracy and, second, by the highest interpretability is selected from that front approximation. The results from all partial experiments are averaged. The experiment is then repeated 10 times for different initializations of our approach. The averaged results are shown in Table 10. The only available alternative approach is the aforementioned decision tree of [24] with 7:3 learning-test data split ratio. Figure 6 summarizes, in an analogous form as in Figure 5, details of our approach (the left part of a block scheme of Figure 6) and the alternative method [24] (the right part of that block scheme) for the purpose of their comparative analysis in terms of the cross-validation-based test-data accuracy and interpretability. The following stages are distinguished in performing that task: data preprocessing, data partitioning, preparation of single *k*-th experiment for 10-fold cross validation (exclusively for our approach), learning process, and calculation of final results. The method of [24] is outperformed by our approach both in terms of system's accuracy and interpretability. Our SPEA3-based approach slightly outperforms its SPEA2-based counterpart in terms of accuracy whereas both of them are characterized by comparable interpretability.

Concluding the experimental section of our work, it is worth emphasizing that measurement of airline passenger satisfaction is a key factor for improving service quality in airline companies [9]. In turn, passenger-satisfaction-based service quality is a strategic tool for gaining competitive advantage [15]. Various specialized companies and agencies such as, e.g., J.D.POWER (see its last report [28]) or American Customer Satisfaction Index (see, e.g., its last ACSI Travel Report [29]) carry out, process, and analyze airline passenger satisfaction surveys to target performance activities that—by attracting more passengers—have a direct impact on profits and reputation.

As far as attributes affecting airline passenger satisfaction are concerned, "In-Flight Wi-Fi Service" and "Simplicity of Online Booking" (an analogical attribute to "Ease of Online Booking" in our research) have been identified in [30] as those which should be optimized by airlines. According to [31], "F&B" (i.e., catering service) and "In-flight entertainment" are principal attributes that affect passenger satisfaction. In turn, four attributes, i.e., "Ease of online booking", "e-ticketing", "Boarding" (analogical attributes to "Online boarding" in our research), and "Clearance time" have been selected in [32] as significant and important factors considered by airline passengers.

In contrast to different methods formulating various sets of attributes affecting airline passenger satisfaction, our approach—employing the modern fuzzy-genetic business-intelligence approach—discovers from the huge set of representative data not only a collection of most important attributes but also formulates a hierarchy of their significance. Moreover, our approach precisely (in percentage values shown in Tables 6–9) determines the level of significance of particular attributes from the airline passenger satisfaction perspective. In order to discover the real significance hierarchy of input attributes, the analysis of the effect of possible "overlapping" of some input attributes over the other ones is also performed. Even more, our approach generates collections of linguistic, fuzzy rules (shown in Tables 4 and 5) that provide a precise and easily-interpretable insight into the mechanisms "connecting" the selected input attributes with airline passenger satisfaction or dissatisfaction. For instance, our approach is able to discover in the airline data not only rather obvious mechanisms (see, e.g., rule No. 5 in solution No. 5: IF Seat comfort is high AND Flight distance is long THEN Passenger is satisfied) but also less obvious ones (see, e.g., rule No. 10 in solution No. 6: IF Inflight WiFi service is low or medium AND Customer type is loyal customer AND Ease of online booking is high AND Leg room service is high THEN Passenger is satisfied). The last rule says that for loyal customers, WiFi services are not significant when online booking and leg room services are of a high quality—perhaps, this rule relates to older passengers who usually do not use WiFi devices. Another such an example is rule No. 9 in solution No. 5: IF Inflight entertainment is high AND Inflight WiFi service is low or medium AND Customer type is disloyal customer AND Flight distance is short THEN Passenger is neutral or dissatisfied. This rule says that high quality of inflight entertainment is not enough to satisfy disloyal passengers travelling on short distances

when WiFi services are low—perhaps, this rule concerns young passengers travelling on short distances and only interested in inflight WiFi services.

5. Conclusions

This paper presents the application of our MOEOA-based knowledge-discovery business-intelligence technique (fuzzy rule-based classification systems) characterized by genetically optimized interpretability-accuracy trade-off to decision support related to airline passenger satisfaction problems. These problems include, first, discovering—in an automatic way—in a large and representative set of data describing airline passenger satisfaction, optimized collections of linguistic, fuzzy classification rules uncovering the "connections" of input attributes with airline passenger satisfaction or dissatisfaction. Second, the considered problems include determining, in a precise and quantitative way, the level of significance (and thus, the formulation of the significance hierarchy) of particular input attributes from the airline passenger satisfaction perspective. Moreover, in order to discover the real significance hierarchy of input attributes, the effect of possible "overlapping" of some of them over the other ones is carefully analyzed.

The main theoretical contribution of our work, in general, consists in introducing our modern MOEOA-based fuzzy-genetic business-intelligence approach with optimized interpretability-accuracy trade-off to broadly understood airline passenger satisfaction decision support. The interpretability and transparency (i.e., the ability to provide the user with compact and understandable explanations and justifications of the decisions proposed) and the accuracy (i.e., the ability to generate precise and correct decisions) are the fundamental aspects of the operation of any decision support systems including those in the aviation industry. On the other hand, compact, linguistic, fuzzy classification rules—due to their easy-to-grasp interpretation and readability—belong to the most effective knowledge-representation schemes in the considered and also in many other domains.

The main experimental contribution our work is twofold. First, it is the application of our approach to recently published and accessible at Kaggle's repository airline passengers satisfaction data set containing 259,760 records. The aspects already listed in the first paragraph of this Conclusions section have been addressed. Second, by means of cross-validation-based experiments, we show that our approach outperforms the alternative method of [24] in terms of both the interpretability and accuracy of the solutions obtained (the paper [24], to our knowledge, is the only available reference addressing the considered and recently published airline passenger satisfaction data set). We also hope that the findings in this research provide insights that could be used by managers and practitioners from aviation industry in defining service strategies and policies that improve airline passenger satisfaction and, consequently, airline reputation and profits.

Our further work will concentrate on two aspects. First, we intend to investigate additional attributes characterizing airline passenger satisfaction. They include "Disembarkation efficiency" pointed out in [9] as the most significant attribute characterizing the flight stage immediately after landing as well as "Announcement of delay and arrival", "Degree of courtesy of staff", and "Adult cost" indicated in [33] as important attributes with which the passengers are most unsatisfied. Second, we intend to concentrate on improving the systems' interpretability-accuracy trade-off optimization, which is essential for generating highly interpretable and accurate modern intelligent decision systems (cf. explainable artificial intelligence [34,35] or interpretable machine learning [36,37]).

Table 6. Illustration of attribute presence and significance in airline passenger satisfaction data.

Attribute Name	$ACC_1^{(tst)}$ / $\Delta ACC_j^{(tst)}, j = 2,3,\dots$	Attribute Presence in the Rules of Solution No.:									
		1	2	3	4	5	6	7	8	9	10
Part A—Number of input attributes: 21											
Inflight entertainment	75.2%	■	■	■	■	■	■	■	■	■	■
Seat comfort	+2.0%		■	■	■	■	■	■	■	■	■
Type of travel	+3.2%				■	■	■	■	■	■	■
Inflight WiFi service					■	■	■	■	■	■	■
Customer type	+2.3%					■	■	■	■	■	■
Flight distance	+1.7%						■	■	■	■	■
Ease of online booking	+1.9%							■	■	■	■
Leg room service								■	■	■	■
On-board service	+1.7%							■	■	■	■
Baggage handling								■	■	■	■
Checkin service								■	■	■	■
Cleanliness	+0.1%								■	■	■
Departure delay in minutes									■	■	■
Class										■	■
Departure/Arrival time conv.	+0.3%									■	■
Gate location										■	■
Food and drink										■	■
Online boarding											■
Arrival delay in minutes	+0.1%										■
Gender											■
Age											■
Part B—Number of input attributes: 20											
Online boarding	71.6%	■	■	■	■	■	■	■	■	■	■
Ease of online booking	+3.7%		■	■	■	■	■	■	■	■	■
On-board service	+1.9%			■	■	■	■	■	■	■	■
Seat comfort	+2.9%				■	■	■	■	■	■	■
Type of travel	+2.3%					■	■	■	■	■	■
Flight distance						■	■	■	■	■	■
Customer type	+2.8%						■	■	■	■	■
Inflight WiFi service	+0.8%							■	■	■	■
Gender								■	■	■	■
Baggage handling	+1.1%								■	■	■
Leg room service										■	■
Cleanliness	+0.5%									■	■
Food and drink										■	■
Arrival delay in minutes											■
Checkin service											■
Age	+0.0%										■
Departure delay in minutes											■
Class											
Departure/Arrival time conv.											
Gate location											
Part C—Number of input attributes: 19											
Ease of online booking	69.0%	■	■	■	■	■	■	■	■	■	■
Seat comfort	+2.6%	■	■	■	■	■	■	■	■	■	■
On-board service	+5.2%			■	■	■	■	■	■	■	■
Type of travel	+3.5%				■	■	■	■	■	■	■
Customer type	+2.4%					■	■	■	■	■	■
Inflight WiFi service	+0.9%						■	■	■	■	■
Flight distance	+1.3%							■	■	■	■
Leg room service								■	■	■	■
Gender	+1.2%								■	■	■
Class	+0.9%									■	■
Baggage handling										■	■
Food and drink										■	■
Arrival delay in minutes	+0.6%										■
Departure/Arrival time conv.											■
Age											■
Cleanliness											
Checkin service											
Departure delay in minutes											
Gate location											

In the left margin of each part, the attribute significance is indicated from *High* (top) to *Low* (bottom), labelled "Attribute significance".

Table 7. (Continuation of Table 6) Illustration of attribute presence and significance in airline passenger satisfaction data.

Vertical axis for each part denotes *Attribute significance* from *High* (top) to *Low* (bottom).

Attribute Name	$ACC_1^{(tst)}$ / $\Delta ACC_j^{(tst)}, j=2,3,\dots$	1	2	3	4	5	6	7	8	9	10
Part D—Number of input attributes: 18											
Inflight WiFi service	66.8%	■	■	■	■	■	■	■	■	■	■
Seat comfort	+2.1%		■	■	■	■	■	■	■	■	■
Class	+6.7%				■	■	■	■	■	■	■
Baggage handling	+2.8%							■	■	■	■
Type of travel	⎫								■	■	■
Customer type	⎬ +2.5%								■	■	■
Checkin service	⎭								■	■	■
Gender	⎫ +0.1%									■	■
Age	⎭									■	■
Departure/Arrival time conv.	⎫										■
Gate location	⎬										■
Leg room service	⎬ +0.0%										■
Arrival delay in minutes	⎭										■
On-board service											
Food and drink											
Cleanliness											
Departure delay in minutes											
Flight distance											
Part E—Number of input attributes: 17											
Leg room service	66.7%	■	■	■	■	■	■	■	■	■	■
On-board service	+3.8%		■	■	■	■	■	■	■	■	■
Seat comfort	+3.3%			■	■	■	■	■	■	■	■
Type of travel	+1.4%					■	■	■	■	■	■
Baggage handling	+1.7%							■	■	■	■
Customer type	+1.5%							■	■	■	■
Gender	+2.8%								■	■	■
Flight distance	+0.0%									■	■
Cleanliness	+0.3%										■
Class											
Age											
Departure/Arrival time conv.											
Checkin service											
Arrival delay in minutes											
Food and drink											
Gate location											
Departure delay in minutes											
Part F—Number of input attributes: 16											
On-board service	66.3%	■		■	■	■	■	■	■	■	■
Seat comfort	⎫ +6.8%		■		■	■	■	■	■	■	■
Cleanliness	⎭				■	■	■	■	■	■	■
Customer type	+2.5%							■	■	■	■
Type of travel	+3.3%								■	■	■
Baggage handling	+0.7%									■	■
Departure/Arrival time conv.	+1.4%										■
Gender											
Flight distance											
Class											
Age											
Checkin service											
Arrival delay in minutes											
Food and drink											
Gate location											
Departure delay in minutes											

Table 8. (Continuation of Table 7) Illustration of attribute presence and significance in airline passenger satisfaction data.

Part G—Number of input attributes: 15

	Attribute Name	$ACC_1^{(tst)}$ $\Delta ACC_j^{(tst)}, j = 2,3,\ldots$	1	2	3	4	5	6	7	8	9	10
High ↑	Seat comfort	64.5%	■	■	■	■	■	■	■	■	■	■
	Baggage handling	+2.7%		■				■	■	■	■	■
	Type of travel	} +2.6%			■	■	■	■	■	■	■	■
	Flight distance				■	■	■	■	■	■	■	■
Attribute sign.	Customer type	+1.6%					■	■	■	■	■	■
	Class	+0.7%								■	■	■
↓	Age	+1.0%									■	■
	Departure/Arrival time conv.	+0.1%										■
	Cleanliness											
	Gender											
	Checkin service											
	Arrival delay in minutes											
	Food and drink											
Low	Gate location											
	Departure delay in minutes											

Attribute Presence in the Rules of Solution No.

Part H—Number of input attributes: 14

	Attribute Name	$ACC_1^{(tst)}$ $\Delta ACC_j^{(tst)}, j = 2,3,\ldots$	1	2	3	4	5	6	7	8	9	10
High ↑	Cleanliness	63.9%	■	■	■	■	■	■	■	■	■	■
	Customer type	} +10.7%		■		■	■	■	■	■	■	■
	Type of travel				■	■	■	■	■	■	■	■
	Gender	+0.4%						■	■	■	■	■
Attribute sign.	Baggage handling	} +1.5%								■	■	■
	Food and drink									■	■	■
	Gate location	+0.5%									■	■
	Age											■
	Class											■
↓	Departure/Arrival time conv.	} +0.5%										■
	Checkin service											■
Low	Arrival delay in minutes											■
	Flight distance											
	Departure delay in minutes											

Part I—Number of input attributes: 13

	Attribute Name	$ACC_1^{(tst)}$ $\Delta ACC_j^{(tst)}, j = 2,3,\ldots$	1	2	3	4	5	6	7	8	9	10
High ↑	Baggage handling	63%	■	■	■	■	■	■	■	■	■	■
	Type of travel	+4.3%		■				■	■	■	■	■
	Class	+5.1%				■	■	■	■	■	■	■
	Flight distance	−0.1%				■	■	■	■	■	■	■
Attr. sign.	Customer type	+1.8%						■	■	■	■	■
	Departure/Arrival time conv.	+0.5%							■	■	■	■
	Gender	+1.8%								■	■	■
	Age	} +0.1%									■	■
↓	Gate location										■	■
	Food and drink	+0.2%										■
Low	Checkin service											
	Arrival delay in minutes											
	Departure delay in minutes											

Table 9. Final hierarchy of attribute significance—comparison of our approach and alternative method of [24].

		Our Approach		Alternative Approach of [24] (Patlolla (2019))	
		Attribute Name	$ACC_1^{(tst)}$	**Attribute Name**	**Importance**
Attribute significance	→ *High*	Inflight entertainment	75.2%	Inflight entertainment	1.0
		Online boarding	71.6%	Class	0.5206
		Ease of online booking	69.0%	Inflight WiFi service	0.4219
		Inflight WiFi service	66.8%	Seat comfort	0.3580
		Leg room service	66.7%	Ease of online booking	0.3333
		On-board service	66.3%	Leg room service	0.2320
		Seat comfort	64.5%	Online boarding	0.2099
	Low ←	Cleanliness	63.9%	Cleanliness	0.1781
		Baggage handling	63.0%	Type of travel	0.1772

Table 10. Results of our approach and comparison with alternative method of [24] (Patlolla (2019)).

Source	Method	Learn-to-Test Ratio	Number of Runs	Average Accuracy Measures for Learning and Test Data			Average Interpretability Measures			
				$ACC^{(lrn)}$	$ACC^{(tst)}$	$\overline{R}$	$\overline{n}_{ATR}$	$\overline{n}_{FS}$	$\overline{n}_{ATR/R}$
Patlolla (2019)	Decision tree	7:3	1	84.0%	84.0%	n/a	n/a	n/a	n/a
This paper	Our approach based on:								
	SPEA2	1:9	10	88.1%	88.0%	16.8	14.3	28.5	3.1
	SPEA3	1:9	10	88.5%	88.3%	17.7	15.5	32.6	3.3

n/a stands for not available.
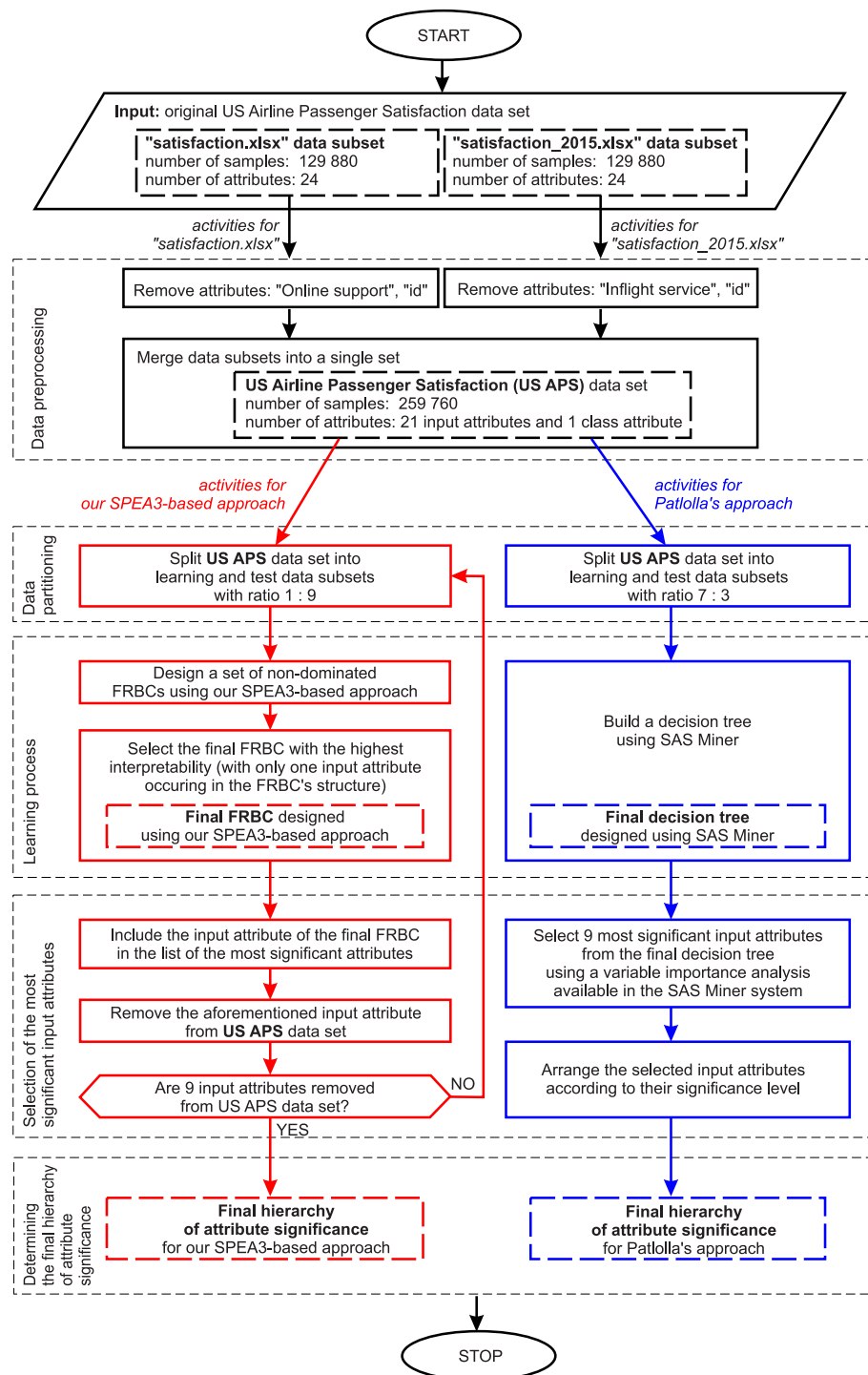
Figure 5. A block scheme for comparative analysis of our approach and alternative Patlolla's method [24] in terms of determining the attribute significance hierarchy.
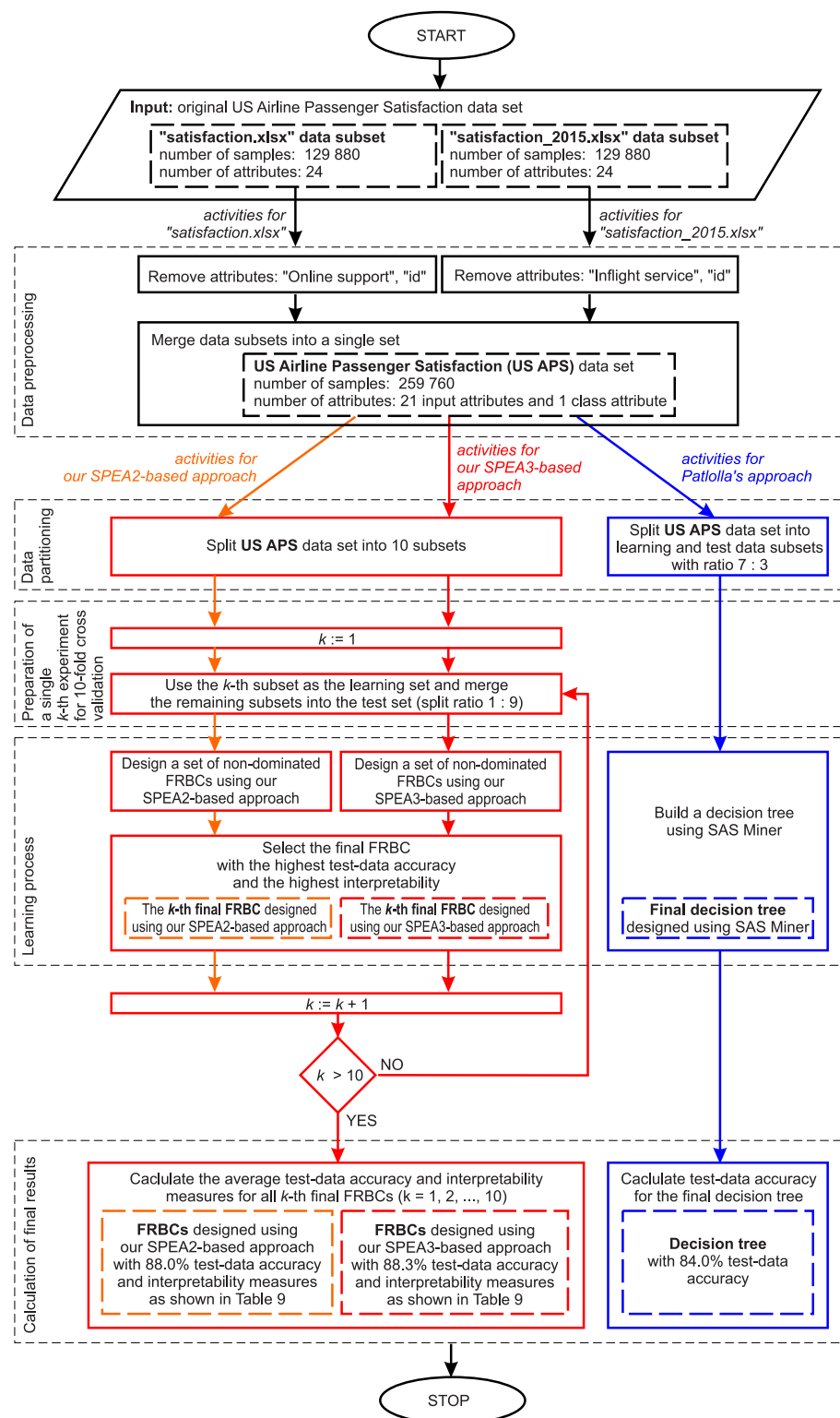
Figure 6. A block scheme for comparative analysis of our approach and alternative Patlolla's method [24] in terms of the cross-validation-based test-data accuracy and interpretability.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sets used in our experiments are accessible at Kaggle's repository—see reference [23].

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

References

1. Grossmann, W.; Rinderle-Ma, S. *Fundamentals of Business Intelligence*; Springer: Berlin/Heidelberg, Germany, 2015.
2. Luhn, H.P. A Business Intelligence System. *IBM J. Res. Dev.* **1958**, *2*, 314–319. [CrossRef]
3. Sharda, R.; Turban, E.; Delen, D.; Aronson, J.; Liang, T. *Business Intelligence and Analytics: Systems for Decision Support*; Always Learning; Pearson: London, UK, 2014.
4. Akerkar, R.; Sajja, P. *Knowledge-Based Systems*, 1st ed.; Jones and Bartlett Publishers, Inc.: Burlington, MA, USA, 2009.
5. Kumar, S.; Zymbler, M. A machine learning approach to analyze customer satisfaction from airline tweets. *J. Big Data* **2019**, *6*. [CrossRef]
6. Bogicevic, V.; Yang, W.; Bujisic, M.; Bilgihan, A. Visual Data Mining: Analysis of Airline Service Quality Attributes. *J. Qual. Assur. Hosp. Tour.* **2017**, *18*, 1–22. [CrossRef]
7. Ban, H.J.; Kim, H.S. Understanding Customer Experience and Satisfaction through Airline Passengers' Online Review. *Sustainability* **2019**, *11*, 4066. [CrossRef]
8. Yakut, I.; Turkoglu, T.; Yakut, F. Understanding Customer's Evaluations Through Mining Airline Reviews. *Int. J. Data Min. Knowl. Manag. Process* **2015**, *5*, 1–11. [CrossRef]
9. Tsafarakis, S.; Kokotas, T.; Pantouvakis, A. A multiple criteria approach for airline passenger satisfaction measurement and service quality improvement. *J. Air Transp. Manag.* **2018**, *68*, 61–75. [CrossRef]
10. Zhang, P.; Fan, C.; Xu, Q.; Ran, X.; Yu, L.; Fang, D.; Zhang, Z. Applications of Business Intelligence Technology in the Airports and Airlines Companies. *Int. J. Appl. Sci. Technol.* **2011**, *1*, 74–78.
11. Rudziński, F. A multi-objective genetic optimization of interpretability-oriented fuzzy rule-based classifiers. *Appl. Soft Comput.* **2016**, *38*, 118–133. [CrossRef]
12. Gorzałczany, M.B.; Rudziński, F. A multi-objective genetic optimization for fast, fuzzy rule-based credit classification with balanced accuracy and interpretability. *Appl. Soft Comput.* **2016**, *40*, 206–220. [CrossRef]
13. Gorzałczany, M.B.; Rudziński, F. Interpretable and accurate medical data classification-a multi-objective genetic-fuzzy optimization approach. *Expert Syst. Appl.* **2017**, *71*, 26–39. [CrossRef]
14. Gorzałczany, M.B.; Rudziński, F. Handling fuzzy systems' accuracy-interpretability trade-off by means of multi-objective evolutionary optimization methods-selected problems. *Bull. Pol. Acad. Sci. Tech. Sci.* **2015**, *63*, 791–798. [CrossRef]
15. Li, W.; Yu, S.; Pei, H.; Zhao, C.; Tian, B. A hybrid approach based on fuzzy AHP and 2-tuple fuzzy linguistic method for evaluation in-flight service quality. *J. Air Transp. Manag.* **2017**, *60*, 49–64. [CrossRef]
16. Fazzolari, M.; Alcala, R.; Nojima, Y.; Ishibuchi, H.; Herrera, F. A review of the application of multiobjective evolutionary fuzzy systems: Current status and further directions. *IEEE Trans. Fuzzy Syst.* **2013**, *21*, 45–65. [CrossRef]
17. Gorzałczany, M.B.; Rudziński, F. Accuracy vs. interpretability of fuzzy rule-based classifiers-an evolutionary approach. In *Artificial Intelligence and Soft Computing-ICAISC 2012*; Lecture Notes in Computer Science; Rutkowski, L., Korytkowski, M., Scherer, R., Tadeusiewicz, R., Zadeh, L.A., Żurada, J.M., Eds.; Springer: Berlin/Heidelberg, Germany, 2012; Volume 7269, pp. 222–230.
18. Gorzałczany, M.B.; Rudziński, F. A modified Pittsburg approach to design a genetic fuzzy rule-based classifier from data. In *Artificial Intelligence and Soft Computing-ICAISC 2010*; Lecture Notes in Computer Science; Rutkowski, L., Scherer, R., Tadeusiewicz, R., Zadeh, L.A., Żurada, J.M., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; Volume 6113, pp. 88–96.
19. Rudziński, F. Finding sets of non-dominated solutions with high spread and well-balanced distribution using generalized strength Pareto evolutionary algorithm. In *Advances in Intelligent Systems Research, Proceedings of the 2015 Conference International Fuzzy Systems Association and European Society for Fuzzy Logic and Technology (IFSA-EUSFLAT-15)*; Alonso, J.M., Bustince, H., Reformat, M., Eds.; Atlantis Press: Gijón, Spain, 2015; Volume 89, pp. 178–185.
20. Gorzałczany, M.B.; Rudziński, F. An improved multi-objective evolutionary optimization of data-mining-based fuzzy decision support systems. In Proceedings of the 2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Vancouver, BC, Canada, 25–29 July 2016; pp. 2227–2234.
21. Gorzałczany, M.B.; Rudziński, F. A multi-objective-genetic-optimization-based data-driven fuzzy classifier for technical applications. In Proceedings of the 2016 IEEE 25th International Symposium on Industrial Electronics (ISIE), Santa Clara, CA, USA, 8–10 June 2016; pp. 78–83.
22. Zitzler, E.; Laumanns, M.; Thiele, L. SPEA2: Improving the strength Pareto evolutionary algorithm for multi-objective optimization. In Proceedings of the Evol. Methods for Design, Optimization and Control with Applications to Industrial Problems, Athens, Greece, 19–21 September 2001; pp. 95–100.
23. John, D. US Airline Passenger Satisfaction Data Set (Version 2). Available online: https://www.kaggle.com/johndddddd/customer-satisfaction (accessed on 26 December 2020).
24. Patlolla, H.R. US Airline Passenger Satisfaction. *MWSUG* **2019**, *RF-079*, 1–12.

25. Gacto, M.J.; Alcala, R.; Herrera, F. Interpretability of linguistic fuzzy rule-based systems: An overview of interpretability measures. *Inf. Sci.* **2011**, *181*, 4340–4360. [CrossRef]

26. Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* **2002**, *6*, 182–197. [CrossRef]

27. Zhuo, X.; Zhang, J.; Son, S.W. Network intrusion detection using word embeddings. In Proceedings of the 2017 IEEE International Conference on Big Data (Big Data), Boston, MA, USA, 11–14 December 2017; pp. 4686–4695.

28. J. D. POWER. North America Airline Satisfaction Study. 2021. Available online: https://www.jdpower.com/business/travel-and-hospitality/north-america-airline-satisfaction-study (accessed on 26 December 2020).

29. American Customer Satisfaction Index (ACSI). ACSI Travel Report 2020–2021. Available online: https://www.theacsi.org (accessed on 26 December 2020).

30. Hayadi, B.H.; Kim, J.M.; Hulliyah, K.; Sukmana, H. Predicting Airline Passenger Satisfaction with Classification Algorithms. *Int. J. Inform. Inf. Syst.* **2021**, *4*, 82–94.

31. Park, S.; Lee, J.S.; Nicolau, J.L. Understanding the dynamics of the quality of airline service attributes: Satisfiers and dissatisfiers. *Tour. Manag.* **2020**, *81*, 104163. [CrossRef]

32. Soomro, D.Y.; Hameed, D.I.; Shakoor, R.; Kaimkhani, S. Factors effecting consumer preferences in airline industry. *Far East J. Psychol. Bus.* **2012**, *7*, 63–72.

33. Abdulsalam, M.A.; Miskeen, B.; Alhodairi, A.M.; Abdullah, R.A.; Ehsan, S. Evaluate the Service Quality of Local Airline Companies in Libya Using Importance-Satisfaction Analysis. *Aust. J. Basic Appl. Sci.* **2013**, *7*, 154–165.

34. Arrieta, A.B.; Diaz-Rodriguez, N.; Del Ser, J.; Bennetot, A.; Tabik, S.; Barbado, A.; Garcia, S.; Gil-Lopez, S.; Molina, D.; Benjamins, R.; et al. Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion* **2020**, *58*, 82–115. [CrossRef]

35. Samek, W.; Montavon, G.; Vedaldi, A.; Hansen, L.; Müller, K.R. (Eds.) *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*; Lecture Notes in Artificial Intelligence; Springer Nature Switzerland AG: Cham, Switzerland, 2019.

36. Monar, C. Interpretable Machine Learning. A Guide for Making Black Box Models Explainable. Published online under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. 2020. pp. 1–318. Available online: https://christophm.github.io/interpretable-ml-book/ (accessed on 26 December 2020).

37. Escalante, H.J.; Escalera, S.; Guyon, I.; Baró, X.; Güçlütürk, Y.; Güçlü, U.; Gerven, M. (Eds.) *Explainable and Interpretable Models in Computer Vision and Machine Learning*; The Springer Series on Challenges in Machine Learning; Springer Nature Switzerland AG: Cham, Switzerland, 2018.