Notes on Intuitionistic Fuzzy Sets
Print ISSN 1310–4926, Online ISSN 2367–8283
2025, Volume 31, Number 4, 496–510
DOI: 10.7546/nifs.2025.31.4.496-510

k-NN intuitionistic fuzzy classifier

Eulalia Szmidt [1,2], **Janusz Kacprzyk**[1,2]
and Paweł Bujnowski [1]

[1] Systems Research Institute, Polish Academy of Sciences
ul. Newelska 6, 01–447 Warsaw, Poland
e-mails: {szmidt,kacprzyk,pbujno}@ibspan.waw.pl

[2] WIT Academy
ul. Newelska 6, 01-447 Warsaw, Poland

Received: 13 October 2025 **Revised:** 11 November 2025
Accepted: 14 November 2025 **Online First:** 27 November 2025

Abstract: The k-NN (k-Nearest Neighbor) classifier is one of the commonly used classifiers. We present a modification of this classifier based on Atanassov's intuitionistic fuzzy sets (IFSs, for short). We show, using benchmark data from *UCI Machine Learning Repository*, that the classifier we propose achieves very good results.

Keywords: Intuitionistic fuzzy sets, Classification, k-NN classifier.
2020 Mathematics Subject Classification: 03E72.

1 Introduction

Atanassov's intuitionistic fuzzy sets [1–3] are a generalization of the fuzzy sets of Zadeh [57]. The IFSs can be viewed as a tool that may help better model the systems in the presence of a lack of knowledge. An advantage of the IFSs is an inherent possibility to take a lack of knowledge into account by using the so-called hesitation margin or intuitionistic fuzzy index.

The k-Nearest Neighbor (k-NN) classifier is one of the simplest yet most widely used methods in pattern recognition and machine learning. Its popularity results from its intuitive principles and effectiveness in many real-world applications.

In past years, many approaches have been proposed to improve the performance and robustness of the k-NN classifier. Among them, fuzzy and intuitionistic fuzzy approaches have proven particularly useful for handling data imprecision and ambiguity (see [7,11,15,19,58]). The very idea of bridging intuitionistic fuzzy sets and the k-NN classifer dates back to the mid 1990s in papers by Hadjitodorov [12–14], Kuncheva [15], and later by Todorova and Vassilev [56].

In this paper, we propose a novel modification of the k-NN classifier that utilizes Atanassov's intuitionistic fuzzy sets (IFSs) to represent input data. Additionally, we introduce a data transformation based on one of the well-known intuitionistic fuzzy operators, which significantly improves the classification performance.

To demonstrate the effectiveness of the proposed approach, we present its results for Quinlan's example (Quinlan [18]) and benchmark datasets from the *UCI Machine Learning Repository* (https://archive.ics.uci.edu/datasets). The obtained results show a clear improvement in both overall classification accuracy and the accuracy of individual class recognition compared to the results achieved by various methods implemented in WEKA.

2 A brief introduction to intuitionistic fuzzy sets

One of the possible generalizations of a fuzzy set in X (Zadeh [57]) given by

$$A' = \{\langle x, \mu_{A'}(x)\rangle | x \in X\}, \tag{1}$$

where $\mu_{A'}(x) \in [0,1]$ is the membership function of the fuzzy set A', is an IFS (see [1–3]) A is given by

$$A = \{\langle x, \mu_A(x), \nu_A(x)\rangle | x \in X\}, \tag{2}$$

where: $\mu_A : X \to [0,1]$ and $\nu_A : X \to [0,1]$ such that

$$0 \le \mu_A(x) + \nu_A(x) \le 1 \tag{3}$$

and $\mu_A(x)$, $\nu_A(x) \in [0,1]$ denote a degree of membership and a degree of non-membership of $x \in A$, respectively. (See Szmidt and Baldwin [21] for assigning memberships and non-memberships for IFSs from data.)

Obviously, each fuzzy set may be represented by the following IFS:

$$A = \{\langle x, \mu_{A'}(x), 1 - \mu_{A'}(x)\rangle | x \in X\}. \tag{4}$$

An additional concept for each IFS in X, that is not only an obvious result of (2) and (3) but which is also relevant for applications, we will call

$$\pi_A(x) = 1 - \mu_A(x) - \nu_A(x) \tag{5}$$

a *hesitation margin* of $x \in A$ which expresses a lack of knowledge of whether x belongs to A or not (cf. Atanassov [2]). It is obvious that $0 \le \pi_A(x) \le 1$, for each $x \in X$.

The hesitation margin turns out to be important while considering the distances (Szmidt and Kacprzyk [22, 23, 26, 33, 36], entropy (Szmidt and Kacprzyk [28, 38]), similarity (Szmidt and

Kacprzyk [37, 50]) for the IFSs, etc., i.e., the measures that play a crucial role in virtually all information processing tasks (Szmidt [20]).

The hesitation margin turns out to be relevant for applications – in image processing (cf. Bustince *et al.* [6]), the classification of imbalanced and overlapping classes (cf. Szmidt and Kukier [53, 54], [55]), the classification applying intuitionistic fuzzy trees (cf. Bujnowski [5]), attribute selection [47,48,51], ranking of alternatives [49,52], multiagent decisions, negotiations, voting, group decision making, etc. (cf. [4, 16, 24, 25, 27, 29, 30, 32, 33, 35, 39, 46]), genetic algorithms [17]. Sometimes the concept of the hesitation margin is just indispensable, for example, for a proper definition of the Hausdorff distance [41] and seeing IFSs like different ones from interval- valued fuzzy sets [45].

In this paper we use the three term representation of the IFSs, i.e., take into account membership values μ, non-membership values ν, and hesitation margins π. The tree term representation is is very useful especially from practical points of view (cf. Szmidt [20], Szmidt and Kacprzyk [26,28,31,34,39–44,46]).

We also use an algorithm [21] of how to derive IFS parameters of a model from relative frequency distributions (histograms). We show the determination of IFS parameters in the case of Quinlan's example.

3 Models of a classifier error

Traditionally *accuracy* of a classifier is measured as the percentage of instances that are correctly classified, and *error* is measured as the percentage of incorrectly classified instances (unseen data). But when the considered classes are imbalanced or when misclassification costs are not equal both the accuracy and the error are not sufficient.

3.1 Confusion matrix

The confusion matrix (Table 1) is often used to assess a two–class classifier.

Table 1. The Confusion Matrix

	Tested Legal	**Tested Illegal**
Actual Legal	TP	FN
Actual Illegal	FP	TN

The meanings of the symbols are:

- TP – the number of correctly classified legal points,
- FN – the number of incorrectly classified legal points as illegals,
- FP – the number of incorrectly classified illegal points as legals,
- TN – the number of correctly classified illegal points,

and

$$ACC = \frac{\text{correctly classified points}}{\text{total points}} = \frac{TP + TN}{TP + TN + FP + FN}; \tag{6}$$

$$TPR = \frac{\text{legals correctly classified}}{\text{total legals}} = \frac{TP}{TP + FN} \tag{7}$$

$$FNR = \frac{\text{legals incorrectly classified}}{\text{total legals}} = 1 - TPR = \frac{FN}{TP + FN} \tag{8}$$

$$FPR = \frac{\text{illegals incorrectly classified}}{\text{total illegals}} = \frac{FP}{FP + TN} \tag{9}$$

$$TNR = \frac{\text{illegals correctly classified}}{\text{total illegals}} = 1 - TPR = \frac{TN}{TN + FP} \tag{10}$$

4 k-Nearest Neighbors algorithm for data expressed via Atanassov's intuitionistic fuzzy sets

The k-Nearest Neighbors classifier (k-NN classifier) is a simple and rather effective way of classifying data. The k-NN classifier determines the class for a new element by comparing it to the elements already seen in the training set. The method looks for the closest k examples, based on a chosen distance measure, and uses them to make the prediction. Many types of distance measures can be applied but the Euclidean distance is the one most commonly used.

We will use the normalized Euclidean distance for IFSs [26]:

$$e_{IFS}(A, B) = \Big(\frac{1}{2n}\sum_{i=1}^{n}\big(\mu_A(x_i) - \mu_B(x_i)\big)^2 + \big(\nu_A(x_i) - \nu_B(x_i)\big)^2 + \big(\pi_A(x_i) - \pi_B(x_i)\big)^2\Big)^{\frac{1}{2}} \tag{11}$$

The values of $e_{IFS}(A, B)$ distance are from the interval $[0, 1]$.

```
Algorithm: k-Nearest Neighbors (k-NN) Algorithm
Input:  Training set D = {(x1, y1), (x2, y2)}, ..., (xn, yn)\};
         Number of neighbors k;
         Distance metric e_{IFS}(..., ...);
         New sample x_new.
Output: Predict label y_new

Compute distance e_{IFS}(i)= e_{IFS}(x_new, xi) for each (xi, yi) in D.

Sort all samples in D by increasing distance e_{IFS}(i).

Select the k nearest samples: N_k = {(xj, yj)} for j = 1, ..., k

Assign x_new to the class c that maximizes
       count of {yj = c | (xj, yj) \in N_k}

Predict label y_new
```

Remark. When the samples are described by several attributes, the distance between individual attributes is computed, these distances are summed, and the mean distance value is then determined.

Modification of the algorithm

The novelty of the above algorithm lies in describing the classified data using intuitionistic fuzzy sets. To obtain better classification results, we make use of the presence of the hesitation margin in the data description and modify the data before applying the above algorithm.

We use the operator D_α [2], which means that an original instance described in terms of intuitionistic fuzzy sets, i.e., $x_i(\mu_i, \nu_i, \pi_i)$, is modified (in the case of a single attribute) in the following way:

$$\mu_i^1 := \mu_i + \alpha \cdot \pi_i, \tag{12}$$

$$\pi_i^1 := (1 - \alpha) \cdot \pi_i, \tag{13}$$

$$\nu_i^1 := \nu_i, \tag{14}$$

where $\alpha \in [0.1]$.

In the experiments we conducted, we apply the modification (12)–(14) to all instances in each of the analyzed datasets. We do this before running the k-NN algorithm. Having the modified data, we randomly divide them into a training part (70%) and a test part (30%). Both the data division and the k-NN procedure are repeated 100 times using random splits for each dataset.

5 Results

We will present results for Quinlan's example and benchmark datasets from the *UCI Repository* to demonstrate the behavior of our approach.

5.1 Results for Quinlan's example

For clarity we use the famous Quinlan's example [18], the so-called "Saturday Morning" in which classification with nominal data is considered. We have objects described by attributes. Each attribute represent a feature and takes on discrete, mutually exclusive values. For example, if the objects were "Saturday Mornings" and the classification involved the weather, possible attributes might be [18]:

- **outlook**, with values {sunny, overcast, rain},
- **temperature**, with values {cold, mild, hot},
- **humidity**, with values {high, normal}, and
- **windy**, with values {true, false},

Altogether, the above attributes provide a zero-order language for characterizing objects in the universe (the attributes are nominal). A particular Saturday morning, an *example*, might be described as: outlook: overcast; temperature: cold; humidity: normal; windy: false.

Each object (example) belongs to one of mutually exclusive classes, C. We assume that there are only two classes, i.e., $C = \{P, N\}$, where: P denotes the set of *positive examples*, and N denotes the set of *negative examples*. There are 14 training examples as shown in Table 2. Each training example e is represented by the attribute-value pairs, i.e., $\{(A_i, a_{i,j}); i = 1, \ldots, l_i\}$ where A_i is an attribute, $a_{i,j}$ is its value – one of possible j values (for each i-th attribute j can be different, e.g., for *outlook*: $j = 3$, for *humidity*: $j = 2$, etc.).

Table 2. The "Saturday Morning" data from [18]

No.	**Attributes**				**Class**
	Outlook	**Temperature**	**Humidity**	**Windy**	
1	sunny	hot	high	false	N
2	sunny	hot	high	true	N
3	overcast	hot	high	false	P
4	rain	mild	high	false	P
5	rain	cool	normal	false	P
6	rain	cool	normal	true	N
7	overcast	cool	normal	true	P
8	sunny	mild	high	false	N
9	sunny	cool	normal	false	P
10	rain	mild	normal	false	P
11	sunny	mild	normal	true	P
12	overcast	mild	high	true	P
13	overcast	hot	normal	false	P
14	rain	mild	high	true	N

First, we make use of frequency description of the problem (see Table 3). In its spirit the method proposed is close to that of De Carvalho *et al.* [8–10] who use histograms to derive some proximity measures.

The frequency measure (Table 3) used for description of the data (Table 2):

$$f(A_i, a_{i,j}, C) = V(C; A_i = a_{i,j})/p_C, \tag{15}$$

where $C = \{P, N\}$; $V(C; A_i = a_{i,j})$ – the number of training examples of C for which $A_i = a_{i,j}$; p_C – the number of the training examples of C.

Table 3. The frequencies obtained

	Outlook			**Temperature**			**Humidity**		**Windy**	
	S	O	R	H	M	C	H	N	T	F
Positive	$2/9$	$4/9$	$3/9$	$2/9$	$4/9$	$3/9$	$3/9$	$6/9$	$3/9$	$6/9$
Negative	$3/5$	$0/5$	$2/5$	$2/5$	$2/5$	$1/5$	$4/5$	$1/5$	$3/5$	$2/5$

To describe and classify the "Saturday Morning" data via the intuitionistic fuzzy sets, we use an algorithm proposed in [21] to assign the parameters of an intuitionistic fuzzy model which describes the attributes (the relative frequency distribution functions given in Table 3 were the starting point of the algorithm). The assigned description of the attributes in terms of intuitionistic fuzzy sets are given in Table 4 and are used for further calculations.

The main idea is to use the values of the membership and non-membership degrees, and of the hesitation margin. So we have a counterpart table in which we have a description of the problem (Table 2) in terms of intuitionistic fuzzy sets (see Tables 4, 5), i.e., $(\mu(.), \nu(.), \pi(.))$; for instance, $(0, 0.33, 0.67)$ is for "sunny".

Table 4. The counterpart intuitionistic fuzzy model

	Outlook			**Temperature**			**Humidity**		**Windy**	
	S	O	R	H	M	C	H	N	T	F
Hesitation margins	0.67	0	0.69	0.67	1	0.49	0.67	0.4	0.67	0.8
Membership values	0	1	0.2	0	0	0.4	0	0.6	0	0.2
Non-membership values	0.33	0	0.11	0.33	0	0.11	0.33	0	0.33	0

Table 5. The "Saturday Morning" data in terms of A-IFSs

No.	**Attributes**				**Class**
	Outlook	**Temperature**	**Humidity**	**Windy**	
1	$(0, 0.33, 0.67)$	$(0, 0.33, 0.67)$	$(0, 0.33, 0.67)$	$(0.2, 0, 0.8)$	N
2	$(0, 0.33, 0.67)$	$(0, 0.33, 0.67)$	$(0, 0.33, 0.67)$	$(0, 0.33, 0.67)$	N
3	$(1, 0, 0)$	$(0, 0.33, 0.67)$	$(0, 0.33, 0.67)$	$(0.2, 0, 0.8)$	P
4	$(0.2, 0.11, 0.69)$	$(0, 0, 1)$	$(0, 0.33, 0.67)$	$(0.2, 0, 0.8)$	P
5	$(0.2, 0.11, 0.69)$	$(0.4, 0.11, 0.49)$	$(0.6, 0, 0.4)$	$(0.2, 0, 0.8)$	P
6	$(0.2, 0.11, 0.69)$	$(0.4, 0.11, 0.49)$	$(0.6, 0, 0.4)$	$(0, 0.33, 0.67)$	N
7	$(1, 0, 0)$	$(0.4, 0.11, 0.49)$	$(0.6, 0, 0.4)$	$(0, 0.33, 0.67)$	P
8	$(0, 0.33, 0.67)$	$(0, 0, 1)$	$(0, 0.33, 0.67)$	$(0.2, 0, 0.8)$	N
9	$(0, 0.33, 0.67)$	$(0.4, 0.11, 0.49)$	$(0.6, 0, 0.4)$	$(0.2, 0, 0.8)$	P
10	$(0.2, 0.11, 0.69)$	$(0, 0, 1)$	$(0.6, 0, 0.4)$	$(0.2, 0, 0.8)$	P
11	$(0, 0.33, 0.67)$	$(0, 0, 1)$	$(0.6, 0, 0.4)$	$(0, 0.33, 0.67)$	P
12	$(1, 0, 0)$	$(0, 0, 1)$	$(0, 0.33, 0.67)$	$(0, 0.33, 0.67)$	P
13	$(1, 0, 0)$	$(0, 0.33, 0.67)$	$(0.6, 0, 0.4)$	$(0.2, 0, 0.8)$	P
14	$(0.2, 0.11, 0.69)$	$(0, 0, 1)$	$(0, 0.33, 0.67)$	$(0, 0.33, 0.67)$	N

By applying the k-NN classifier described above, which represents instances using intuitionistic fuzzy sets, we obtained the results shown in the first row of Table 6, corresponding to the parameter $\alpha = 0$, that is, the results from WEKA.

From WEKA we obtained the best result for function.Logistic: accuracy $= 66\%$, TPR $= 71\%$, TNR $= 58\%$. However, for the counterpart from WEKA of k-NN classifier ($k = 3$), i.e., for lazy.

IBk ($k = 3$) we obtained the following results: accuracy $= 62\%$, TPR $= 77\%$ (bigger class), TNR $= 40\%$ (smaller class).

For the new classifier presented here with the parameter $\alpha = 0$, the accuracy $= 74\%$, TPR $= 65.3\%$ (smaller class), TNR $= 79\%$ (bigger class). After applying the modification (12)–(14), the results improved significantly (see Table 6).

Table 6. Results for dataset 'Weather'

α	ACC	TPR	FNR	FPR	TNR
0	74.4	65.3	34.7	20.6	79.4
0.1	79.5	71.8	28.2	16.3	83.7
0.2	86.9	75.6	24.4	6.9	93.1
0.4	96.3	89.9	10.0	0.16	99.8
0.6	99.6	98.7	1.3	0	100
0.8	99.4	98.3	1.7	0	100
1	99.7	99.0	0.97	0	100

5.2 Results for dataset 'Heart'

https://archive.ics.uci.edu/dataset/95/spect+heart

Number of instances: 267 (55+212)

Number of attributes: 22 features (binary) + 1 target

Task type: Classification - predicting the presence of heart disease

The best classification results for the 'Heart' dataset obtained from WEKA were achieved with the LMT (Logistic Model for Trees). The accuracy is 84%, but the model shows poor recognition of patients with heart disease, with a TPR of only 54%.

The proposed k-NN classifier, using Intuitionistic fuzzy sets achieves 100% accuracy with the parameter value $\alpha = 0.2$ (see Table 7).

Table 7. Results for dataset 'Heart'

α	ACC	TPR	FNR	FPR	TNR
0	77.6	38.1	61.9	11.9	88.1
0.1	99.5	97.4	2.6	0.0	100.0
0.2	100.0	100.0	0.0	0.0	100.0

5.3 Results for dataset 'Hepatitis'

https://archive.ics.uci.edu/dataset/46/hepatitis

We use the official cleaned version of the data (cases without missing values).

Number of instances: 80 (13 + 67).

Number of attributes: 19 (excluding the class).

Task type: Classification (binary: live vs die).

Classes Die: 13, Live: 67.

The best accuracy as indicated by WEKA is achieved by the Trees Random Forest classifier, with an accuracy 88.5% and a TPR $= 46\%$.

The k-NN classifier proposed in this paper, with the parameter $\alpha = 0.2$, produces excellent results (see Table 8).

Table 8. Results for dataset 'Hepatitis'

α	ACC	TPR	FNR	FPR	TNR
0	85.3	53.3	46.7	8.0	92.0
0.1	99.9	99.9	0.1	0.01	100.0
0.2	100.0	100.0	0.0	0.0	100.0

5.4 Results for dataset 'Ionosphere'

https://archive.ics.uci.edu/ml/datasets/ionosphere

'Ionosphere' is a dataset with two classes (class proportion: 225:126), 351 instances, and 35 attributes. The 35th attribute points out to a class which can be "good" or "bad" according to radar returns from the ionosphere.

The best results among the algorithms in WEKA are achieved by Random Forest, with an accuracy of 93.6%. The detection rate of the minority class in this case is 87.3%.

The k-NN classifier proposed by us achieves better results for the parameter $\alpha = 0.2$ (see Table 9).

Table 9. Results for dataset 'Ionosphere'

α	ACC	TPR	FNR	FPR	TNR
0	90.3	77.4	22.6	2.2	97.7
0.1	97.2	94.5	5.5	1.2	98.8
0.2	99.9	99.9	0.1	0.01	99.9

5.5 Results for dataset 'Diabetes Pima'

https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database

Number of instances: 768.

Number of attributes: 8 input + 1 class (so 9 total).

Classification: diabetes vs no diabetes.

Diabetic cases: 268, Non-diabetic cases: 500.

Among the various classifiers available in WEKA, the Functions.Logistic algorithm achieves the highest accuracy of 77.5%, with a True Positive Rate (TPR) for the diabetes class of 57.1%.

The proposed k-NN classifier, employing the framework of intuitionistic fuzzy sets, achieves perfect patient classification for the parameter $\alpha = 0.4$ (see Table 10).

Table 10. Results for dataset 'Diabetes Pima'

α	ACC	TPR	FNR	FPR	TNR
0	71.9	53.9	46.1	18.4	81.6
0.1	95.8	96.6	3.4	4.7	95.3
0.2	99.3	100.0	0.0	1.1	98.9
0.4	100.0	100.0	0.0	0.0	100.0

5.6 Results for dataset 'Sonar

https://archive.ics.uci.edu/dataset/151/connectionist+bench+sonar+mines+vs+rocks

Number of instances: 208.

Number of attributes: 60 + 1 class attribute.

Classification – distinguishing between sonar signals bounced off a metal cylinder (mine) vs a rock.

Mines (metal cylinder): 111 instances, Rocks: 97 instances.

For the Sonar dataset, the best accuracy was obtained using the lazy.IBk classifier from WEKA, reaching 86.2%, with a True Positive Rate (TPR) of 82%.

The proposed k-NN classifier achieves an accuracy of error-free for the parameter $\alpha = 0.4$ (see Table 11).

Table 11. Results for dataset 'Sonar'

α	ACC	TPR	FNR	FPR	TNR
0	82.6	71.7	28.3	7.3	92.7
0.1	97.7	99.7	0.3	4.0	96.0
0.2	99.9	100.0	0.0	0.1	99.9
0.4	100.0	100.0	0.0	0.0	100.0

6 Conclusions

We proposed a novel k–NN classifier operating on data represented in the form of intuitionistic fuzzy sets. In addition, we introduced a modification of the data based on one of the well-known operators, which resulted in a significant improvement in the classifier's performance, both in terms of overall accuracy and the accuracy of recognizing individual classes.

References

[1] Atanassov, K. (1983). Intuitionistic fuzzy sets. *VII ITKR Session*. Sofia (Centr. Sci.-Techn. Libr. of Bulg. Acad. of Sci., 1697/84) (in Bulgarian).

[2] Atanassov, K. T. (1999). *Intuitionistic Fuzzy Sets: Theory and Applications*. Springer-Verlag, Heidelberg.

[3] Atanassov, K. T. (2012). *On Intuitionistic Fuzzy Sets Theory*. Springer-Verlag, Berlin.

[4] Atanassova, V. (2004). Strategies for decision making in the conditions of intuitionistic fuzziness. *Proceedings of International Conference 8th Fuzzy Days*, Dortmund, Germany, 263–269.

[5] Bujnowski, P., Szmidt, E., & Kacprzyk, J. (2014). Intuitionistic fuzzy decision trees - A new approach. In: *Rutkowski, L., Korytkowski, M., Scherer, R., Tadeusiewicz, R., Zadeh, L., & Zurada, J. (Eds.). Artificial Intelligence and Soft Computing, Part I*. Springer, Switzerland, 181–192.

[6] Bustince, H., Mohedano, V., Barrenechea, E., & Pagola, M. (2006). An algorithm for calculating the threshold of an image representing uncertainty through A-IFSs. *IPMU'2006*, 2383–2390.

[7] Campagner, A., Dorigatti, V., & Ciucci, D. (2020). Entropy-based shadowed set approximation of intuitionistic fuzzy sets. *International Journal of Intelligent Systems*, 35(12), 2118–2139.

[8] De Carvalho, F. A. T. (1994). Proximity coefficients between Boolean symbolic objects. In: *Diday et al. (Eds.). New Approaches in Classification and Data Analysis*. Springer-Verlag, Heidelberg, 387–394.

[9] De Carvalho, F. A. T. (1998). Extension based proximities between Boolean symbolic objects, In: *Hayashi, C. et al. (Eds.), Data Science, Classification and Related Methods*. Springer-Verlag, Tokyo, 370–378.

[10] De Carvalho, F. A. T., & Souza, R. M. C. (1998). Statistical proximity functions of Boolean symbolic objects based on histograms. In: *Rizzi, A., Vichi, M., Bock, H.H. (Eds.). Advances in Data Science and Classification*, Springer-Verlag, Heidelberg, 391–396.

[11] Derrac, J., Garcia, S., & Herrera, F. (2014). Fuzzy nearest neighbor algorithms: Taxonomy, experimental analysis and prospects. *Information Sciences*, 260, 98–119.

[12] Hadjitodorov, S. (1995). An intuitionistic fuzzy sets application to the K-NN method. *Notes on Intuitionistic Fuzzy Sets*, 1(1), 66–69.

[13] Hadjitodorov, S. (1998). Intuitionistic fuzzy versions of K-NN method and their application to respiratory distress syndrome detection. *Notes on Intuitionistic Fuzzy Sets*, 4(4), 62–67.

[14] Hadjitodorov, S. (2000). A F-operator intuitionistic fuzzy version of the nearest neighbor classifier. *Notes on Intuitionistic Fuzzy Sets*, 6(1), 1–6.

[15] Kuncheva, L. (1995). An intuitionistic fuzzy k-nearest neighbors rule. *Notes on Intuitionistic Fuzzy Sets*, 1(1), 56–60.

[16] Maggiora, G., & Szmidt, E. (2021). An intuitionistic fuzzy set analysis of drug-target interactions. *MATCH Communications in Mathematical and in Computer Chemistry*, 85(3), 465–498.

[17] Roeva, O., & Michalikova, A. (2013). Generalized net model of intuitionistic fuzzy logic control of genetic algorithm parameters. *Notes on Intuitionistic Fuzzy Sets*, 19(2), 71–76.

[18] Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1, 81–106.

[19] Syriopoulos, P. K., Kalampalikis, N. G., Kotsiantis, S. B., & Vrahatis, M. N. (2025). kNN Classification: A review. *Annals of Mathematics and Artificial Intelligence*, 93, 43–75.

[20] Szmidt, E. (2014). *Distances and Similarities in Intuitionistic Fuzzy Sets*. Springer.

[21] Szmidt, E., & Baldwin, J. (2006). Intuitionistic fuzzy set functions, mass assignment theory, possibility theory and histograms. *Proceedings of 2006 IEEE World Congress on Computational Intelligence*, pp. 237–243.

[22] Szmidt, E., & Kacprzyk, J. (1996). Group decision making via intuitionistic fuzzy sets. *Proceedings of FUBEST'96*, Sofia, Bulgaria, pp. 107–112.

[23] Szmidt, E., & Kacprzyk, J. (1997). On measuring distances between intuitionistic fuzzy sets. *Notes on Intuitionistic Fuzzy Sets*, 3(4), 1–13.

[24] Szmidt, E., & Kacprzyk, J. (1998). Group decision making under intuitionistic fuzzy preference relations. *Proceedings of IPMU'98*, pp. 172–178.

[25] Szmidt, E., & Kacprzyk, J. (1998). Applications of intuitionistic fuzzy sets in decision making. *Proceedings of EUSFLAT'99*, Navarra, pp. 150–158.

[26] Szmidt, E., & Kacprzyk, J. (2000). Distances between intuitionistic fuzzy sets. *Fuzzy Sets and Systems*, 114(3), 505–518.

[27] Szmidt, E., & Kacprzyk, J. (2001). Analysis of consensus under intuitionistic fuzzy preferences. *Proceedings of International Conference in Fuzzy Logic and Technology*. Leicester, UK, pp. 79–82.

[28] Szmidt, E., & Kacprzyk, J. (2001). Entropy for intuitionistic fuzzy sets. *Fuzzy Sets and Systems*, 118(3), 467–477.

[29] Szmidt, E., & Kacprzyk, J. (2002). Analysis of agreement in a group of experts via distances between intuitionistic fuzzy preferences. *Proceedings of the 9th International Conference IPMU'2002*, Annecy, France, pp. 1859–1865.

[30] Szmidt, E., & Kacprzyk, J. (2002). Evaluation of agreement in a group of experts via distances between intuitionistic fuzzy sets. *Proceedings of 2002 International IEEE Symposium on Intelligent Systems (IS'2002)*, Varna, Bulgaria, pp. 166–170.

[31] Szmidt, E., & Kacprzyk, J. (2004). A concept of similarity for intuitionistic fuzzy sets and its use in group decision making. *Proceedings of 2004 IEEE Conference on Fuzzy Systems*, Budapest, pp. 1129–1134.

[32] Szmidt, E., & Kacprzyk, J. (2005). A new concept of a similarity measure for intuitionistic fuzzy sets and its use in group decision making. In: *Torra, V., Narukawa, Y., & Miyamoto. S. (Eds.). Modelling Decisions for Artificial Intelligence.* Lecture Notes on Artifiial Intelligence, Vol. 3558, pp. 272–282. Springer.

[33] Szmidt, E., & Kacprzyk, J. (2005). Distances between intuitionistic fuzzy sets and their applications in reasoning. In: *Halgamuge, S., & Wang, L. (Eds.). Computational Intelligence for Modelling and Prediction.* Studies in Computational Intelligence, Vol 2. pp. 101–116. Springer.

[34] Szmidt, E., & Kacprzyk, J. (2005). New measures of entropy for intuitionistic fuzzy sets. *Notes on Intuitionistic Fuzzy Sets*, 11(2), 12–20.

[35] Szmidt, E., & Kacprzyk, J. (2006). An application of intuitionistic fuzzy set similarity measures to a multi-criteria decision making problem. *Proceedings of ICAISC 2006*. Lecture Notes on Artificial Intelligence, Vol. 4029, pp. 314–323. Springer-Verlag.

[36] Szmidt, E., & Kacprzyk, J. (2006). Distances between intuitionistic fuzzy sets: Straightforward approaches may not work. *Proceedings of 3rd International IEEE Conference Intelligent Systems (IEEE IS'06)*, London, pp. 716–721.

[37] Szmidt, E., & Kacprzyk, J. (2007). A new similarity measure for intuitionistic fuzzy sets: Straightforward approaches may not work. *Proceedings of 2007 IEEE International Conference on Fuzzy Systems*, pp. 481–486.

[38] Szmidt, E., & Kacprzyk, J. (2007). Some problems with entropy measures for the Atanassov intuitionistic fuzzy sets. *Applications of Fuzzy Sets Theory.* Lecture Notes on Artificial Intelligence, Vol. 4578, pp. 291–297. Springer-Verlag.

[39] Szmidt, E., & Kacprzyk, J. (2009). Ranking of intuitionistic fuzzy alternatives in a multi-criteria decision making problem. In: *Proceedings of NAFIPS 2009*, Cincinnati, USA, June 14–17, 2009, pp. 1-6, doi: 10.1109/NAFIPS.2009.5156417.

[40] Szmidt, E., & Kacprzyk, J. (2010). Dealing with typical values via Atanassov's intuitionistic fuzzy sets. *International Journal of General Systems*, 39(5), 489–506.

[41] Szmidt, E., & Kacprzyk, J. (2011). Intuitionistic fuzzy sets – Two and three term representations in the context of a Hausdorff distance. *Acta Universitatis Matthiae Belii, Series Mathematics*, 19, 53–62.

[42] Szmidt, E., & Kacprzyk, J. (2012). A new approach to Principal Component Analysis for intuitionistic fuzzy data sets. *Greco, S., et al. (Eds.). IPMU 2012*, Part II, CCIS 298, pp. 529–538, Springer-Verlag Berlin Heidelberg.

[43] Szmidt, E., & Kacprzyk, J. (2015). Two and three term representations of intuitionistic fuzzy sets: Some conceptual and analytic aspects. *Proceedings of 2015 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2015)*, pp. 1–8.

[44] Szmidt, E., & Kacprzyk, J. (2017). A Perspective on Differences Between Atanassov's Intuitionistic Fuzzy Sets and Interval-Valued Fuzzy Sets. In: *Torra, V., Dahlbom, V., Narukawa, Y. (Eds.). Fuzzy Sets, Rough Sets, Multisets and Clustering*. Studies in Computational Intelligence, Volume 671, pp. 221–237, Springer.

[45] Szmidt, E., & Kacprzyk, J. (2022). Atanassov's intuitionistic fuzzy sets demystified. In: *Ciucci, D., Couso, I., Medina, J., Slezak, D., Petturiti, D., Bouchon-Meunier, B., & Yager, R. R. (Eds.). Information Processing and Management of Uncertainty in Knowledge-Based Systems. 19th International Conference, IPMU 2022*, Milan, Italy, Proceedings, Part I, pp. 517–527. Communications in Computer and Information Science 1601, Springer.

[46] Szmidt, E., Kacprzyk, J., Atanassova, V., & Bujnowski, P. (2024). Intuitionistic fuzzy sets in group decision making – A novel approach. *Notes on Intuitionistic Fuzzy Sets*, 30(2), 101–112.

[47] Szmidt, E., Kacprzyk, J., & Bujnowski, P. (2020). Attribute selection for sets of data expressed by intuitionistic fuzzy sets. *Proceedings of 2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, Glasgow, UK, pp. 1–7.

[48] Szmidt, E., Kacprzyk, J., & Bujnowski, P. (2021). Three term attribute description of Atanassov's intuitionistic fuzzy sets as a basis of attribute selection. *Proceedings of 2021 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pp. 1–6.

[49] Szmidt, E., Kacprzyk, J., & Bujnowski, P. (2022). Ranking of alternatives described by Atanassov's intuitionistic fuzzy sets – A critical review. *Proceedings of the 2022 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pp. 1–7.

[50] Szmidt, E., Kacprzyk, J., & Bujnowski, P. (2022). Similarity measures for Atanassov's intuitionistic fuzzy sets: Some dilemmas and challenges. *Control and Cybernetics*, 51(2), 249–266.

[51] Szmidt, E., Kacprzyk, J., & Bujnowski, P. (2023). Identification of a sufficient number of the best attributes in the intuitionistic fuzzy models. *Notes on Intuitionistic Fuzzy Sets*, 29(2), 144–156.

[52] Szmidt, E., Kacprzyk, J., Bujnowski, P., Starczewski, J., & Siwocha, A. (2024). Ranking of alternatives described by Atanassov's intuitionistic fuzzy sets – Reconciling some misunderstandings. *Journal of Artificial Intelligence and Soft Computing Research*, 14(3), 237–250.

[53] Szmidt, E., & Kukier, M. (2006). Classification of imbalanced and overlapping classes using intuitionistic fuzzy sets. *Proceedings of 3rd International IEEE Conference on Intelligent Systems IEEE IS'06*, London, 722–727.

[54] Szmidt, E., & Kukier, M. (2008). A new approach to classification of imbalanced classes via Atanassov's intuitionistic fuzzy sets. In: *Wang, H.-S. (Ed.). Intelligent Data Analysis: Developing New Methodologies Through Pattern Discovery and Recovery*. Idea Group, 85–101.

[55] Szmidt, E., & Kukier, M. (2008). Atanassov's intuitionistic fuzzy sets in classification of imbalanced and overlapping classes. In: *Chountas, P., Petrounias, I., & Kacprzyk, J. (Eds.). Intelligent Techniques and Tools for Novel System Architectures.* Springer, Series: Studies in Computational Intelligence, Berlin, Heidelberg, pp. 455–471.

[56] Todorova, L., & Vassilev, P. (2009). Application of K-Nearest Neighbor rule in the case of intuitionistic fuzzy sets for pattern recognition. *Bioautomation*, 13(4), 265–270.

[57] Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8(3), 338–353.

[58] Zeraatkar, S., & Afsari. F. (2021). Interval-valued fuzzy and intuitionistic fuzzy-KNN for imbalanced data classification. *Expert Systems with Applications*, 184, 115–510.