Fuzzy Linguistic Summaries for Hidden Markov Models

Katarzyna Kaczmarek-Majer[1][0000−0003−0422−9366], Michał Baczyński[2][0000−0002−4442−2112], Olgierd Hryniewicz[1][0000−0001−9877−508X], Katarzyna Miś[2][0000−0002−0375−6546], Weronika Mucha[3], and Filip Wichrowski[1]

[1] Systems Research Institute Polish Academy of Sciences,
Newelska 6, 01-147 Warsaw, Poland
k.kaczmarek,f.wichrowski@ibspan.waw.pl

[2] Faculty of Science and Technology, University of Silesia in Katowice,
Bankowa 14, 40-007 Katowice, Poland
michal.baczynski,katarzyna.mis@us.edu.pl

[3] Faculty of Mathematics and Information Science, Warsaw University of Technology
Koszykowa 75, 00-662 Warsaw, Poland
weronika.mucha2.stud@pw.edu.pl

Abstract. Linguistic summaries are an intuitive tool for obtaining analysis and data mining results that are easy to use, even for novice users. Until now, linguistic summarization has been used primarily to describe and facilitate the interpretation of large data sets. This work aims to develop methods enabling the construction of linguistically quantified sentences reflecting both the sequence of observations of a time series as well as the estimated parameters of hidden Markov models. The resulting fuzzy linguistic summaries with hidden Markov models (HMMs) may be exemplified as follows: *"For most observations around 1.1, we have a high exact match rate"*. Preliminary results illustrate the effectiveness of the proposed approach using simulation methods.

Keywords: Fuzzy Linguistic Summaries · Hidden Markov Modelling · Explainability · Data Stream Analysis

1 Introduction

Explainable algorithms enable the comprehension of the datasets and reasoning underlying the predictions that they produce. A truly explainable model should not leave the explanation generation to the users as different explanations may be deduced depending on the background users' knowledge. Despite the fast growth of Explainable Artificial Intelligence (XAI) in recent years, there is still a need for more general, unified theories that approximate the structure and intent of an explanation [2]. Explainable methods shall meet the interpretability, completeness, and quantification criteria [4].

Overall, one can distinguish between methods that explain the black box post-hoc and methods that are, on the other hand, explainable by design. The majority of the explainable by-design algorithms are focused on solving the classification or regression problem, e.g., decision trees and rule-based classifiers. On the other hand, one of the common post-hoc methods is Shapley Additive Explanations [11] based initially on the well-known Shapley values. However, it is often difficult in practice to interpret the resulting graphical plots into simpler terms for a non-technical audience such as domain experts, decision-makers, or end-users [9].

In our previous work [9], we confirmed experimentally that linguistic summaries represent human-consistent information granules and improve the overall explainability of selected classifiers (XGBoost, Random Forest, etc.). We now investigate the structure of a particular class of predictive models, the hidden Markov models (HMMs), and build sentences in natural language that best explain the reasoning and final results. In particular, we aim at the construction of linguistic summaries that may be exemplified as follows: *"For most observations around 1.1, we have a high exact match rate"*.

It needs to be noted that HMMs enable us to consider the temporal structure of data and, thus, are often more adequate than static classifiers. We focus on time series data and model the evolving nature of the observed processes and states. Next, we solve the linguistic summarization problem, which belongs to the class of data-to-text data-mining approaches. Specifically, we consider the semi-continuous hidden Markov models [5] and the Baum-Welch algorithm. In our approach, we consider semi-continuous output probability density functions (pdfs) mainly for computational purposes. The performance of the proposed approach is validated experimentally with simulated time series. The rest of the paper is organized as follows. Section 2 explains the semi-continuous hidden Markov models. Section 3 focuses on the methodology related to fuzzy linguistic summarization. Section 4 presents the proposed approach for linguistic summarization of HMMs. The results of illustrative experiments based on simulated data are presented in Section 5. The conclusions and future directions are discussed in Section 6.

2 Hidden Markov Models for Time Series

Let X_t be a Markov process with the state space $S = s_1, s_2, \ldots, s_N$, and q_t a hidden state in time t. Let Y_t be a stochastic process with a state space of $V = v_1, v_2, \ldots, v_M$, and $\mathcal{O}_t$ state (observation) at time t. The probability of transitioning between these states in consecutive time steps is given by a transition matrix $\mathbf{A} = \{a_{ij}\}$, where

$$a_{ij} = \mathbb{P}(X_t = s_i \mid X_{t-1} = s_j); \quad i, j = 1, \ldots, N. \tag{1}$$

The initial state has a prior probability distribution $\pi = \{\pi_i\}$, where

$$\pi_i = \mathbb{P}(X_1 = s_i); \quad i = 1, \ldots, N. \tag{2}$$

An observation, Y_t, is available at each time step and associated with X_t. In particular, the observation Y_t given X_t is conditionally independent on any other state.

In the case of discrete observations, the model requires the probability of emitting an observation in a given state is given by an emission matrix $\mathbf{B} = \{b_j(k)\}$ where

$$b_j(k) = \mathbb{P}(Y_t = v_k \mid X_t = s_j). \tag{3}$$

HMM can be, therefore, described as a triple

$$\lambda = (\mathbf{A}, \mathbf{B}, \pi). \tag{4}$$

There are several ways to estimate the parameters of a model given a sequence of observations; see, e.g., [13]. The following assumptions are often made regarding the HMMs with discrete observations:

(i) the initial distribution, π, follows a discrete uniform distribution;
(ii) the elements of transition matrix $\mathbf{A}$ are estimated using maximum likelihood approach; and
(iii) the conditional probability density functions (pdfs) of the observations follow a univariate normal distribution, parametrized by the hidden state, e.g., $b_j(Y_k) \sim \mathcal{N}(\mu_{s_j}, \Sigma_{s_j})$.

In the case of continuous observations, one can assume the quasi-probability of emissions modelled using a mixture of normal distributions

$$b_j(\mathcal{O}) = \sum_{m=1}^{M} c_{jm} \mathcal{N}(\mathcal{O}, \boldsymbol{\mu}_{jm}, \Sigma_{jm}). \tag{5}$$

However, a large number of free parameters and a potentially small number of sample data are often problematic. Thus, in this work, we adapt the semi-continuous hidden Markov model [5], also called tied-mixture (TMHMM). It is assumed that in each of the states, we have the same normal distributions, differing only in weight coefficients

$$b_j^*(\mathcal{O}) = \sum_{m=1}^{M} c_{jm} \mathcal{N}(\mathcal{O}, \boldsymbol{\mu}_m, \Sigma_m). \tag{6}$$

This approach allows for a significant reduction in the number of model parameters. In the following, we will consider the normalized probability density function

$$b_j(\mathcal{O}) \coloneqq \frac{b_j^*(\mathcal{O})}{\sum_{j=1}^{N} b_j^*(\mathcal{O})} = \frac{\sum_{m=1}^{M} c_{jm} \mathcal{N}(\mathcal{O}, \boldsymbol{\mu}_m, \Sigma_m)}{\sum_{j=1}^{N} \sum_{m=1}^{M} c_{jm} \mathcal{N}(\mathcal{O}, \boldsymbol{\mu}_m, \Sigma_m)}. \tag{7}$$

Finally, the Baum-Welch reestimation algorithm [3] is used iteratively for the estimation of parameters in Eq. (6). The amount of training data required, as well as the computational complexity of the semi-continuous hidden Markov model, can be significantly reduced in comparison with the continuous mixture hidden Markov model. For further descriptions, see e.g., [5].

3 Fuzzy Linguistic Summaries

3.1 Preliminaries

Linguistic summaries are sentences in natural language that describe numerical data. The first notion of linguistically quantified sentences was introduced in the 80's [14], and since then, linguistic summaries have been confirmed as a powerful tool in various domains, see, e.g., [1, 6–8, 10, 15].

Definition 1 (cf. [14]). *Let* $Y = \{y_1, \ldots, y_n\}$ *be a set of objects and* $A = \{A_1, \ldots, A_m\}$ *a set of attributes characterizing objects from* Y*. A summary of a data set consists of*

– *a summarizer* P*,*
– *a qualifier* R *(optionally),*
– *a quantity in agreement* Q*,*
– *a measure of validity or truth of the summary* T*.*

Example 1. Let us consider simple examples of linguistic summaries:

(i) 'Most of employees earn low salaries.' ('Q y's are P');
(ii) 'Most of young employees earn low salaries.' ('Q R y's are P').

The linguistic terms can play the roles of a summarizer (*low*), a qualifier (*young*) or a linguistic quantifier (*most*).

3.2 Fuzzy linguistic summary

We now further define assumptions and functions comprising the linguistic summaries based on the extended protoforms of the following form:

$$Q\ R\ y\text{'s are } P. \tag{8}$$

Firstly, for the given set of objects Y and set of attributes $A = \{A_1, \ldots, A_m\}$, each attribute A_i is a function $A_i \colon Y \to X_i,\ i = 1, \ldots, m$, where $\emptyset \neq X_i \subset \mathbb{R}$. The set $L(Y, A_i)$ of linguistic terms for $i \in \{1, \ldots, m\}$ is given as follows $L(Y, A_i) = \{l_1^{A_i}, \ldots, l_{k_i}^{A_i}\}$ and it enables the formulation of summaries in natural language. We use type-I fuzzy sets to describe linguistic terms and quantifiers. Also, let $k_i = |L(Y, A_i)|$ for given $i \in \{1, \ldots, m\}$. Furthermore, let $\mathcal{V}$ be the following family of functions:

$$\mathcal{V} = \{V_{i,k} \mid V_{i,k} \colon A_i(Y) \to [0, 1], i = 1, \ldots, m,\ k = 1, \ldots, k_i\}. \tag{9}$$

Now, a fuzzy linguistic summary is given as follows.

Definition 2 (cf. [12]). *Let* $S = (Y, A, \mathcal{P}, \mathcal{R}, Q)$ *be a 5-tuple (quintuple) such that*

(i) Y *is a set of objects,*

(ii) *A is a set of attributes,*

(iii) *$\mathcal{P} \in \mathcal{V}$ is a summarizer, $\mathcal{R} \in \mathcal{V}$ is qualifier, and they satisfy*

$$\exists_{i\in\{1,\ldots,m\}}\exists_{k\in\{1,\ldots,k_i\}}\ V_{i,k} = \mathcal{P} \Rightarrow \forall_{k\in\{1,\ldots,k_i\}}\ V_{i,k} \neq \mathcal{R}, \tag{10}$$

and

$$\exists_{i\in\{1,\ldots,m\}}\exists_{k\in\{1,\ldots,k_i\}}\ V_{i,k} = \mathcal{R} \Rightarrow \forall_{k\in\{1,\ldots,k_i\}}\ V_{i,k} \neq \mathcal{P}, \tag{11}$$

(iv) *$Q\colon B \to [0,1]$ is a linguistic quantifier and $B \in \{\mathbb{R}^+, [0,1]\}$.*

Then S is called a fuzzy linguistic summary.

Note, the condition (*iii*) means that if any element $\mathcal{V}$ plays the role of a summarizer, then any other element of $\mathcal{V}$ which describes the same attribute cannot be a qualifier and vice versa. Linguistic quantifier $Q\colon X \to [0,1]$ can be absolute (e.g. *at most ten*) or proportional (e.g. *about half*); see [14]. This distinction implies that they can have different domains ($\mathbb{R}_+$ for absolute, $[0,1]$ for proportional).

Clearly, for different sets of objects, the informativeness of fuzzy (linguistic) summaries is on various levels. It can be measured by numerous criteria, for instance, by a degree of truth, a degree of imprecision, a degree of covering, and a degree of appropriateness [8]. We use the following notion to assess the quality of fuzzy linguistic summaries.

Definition 3. *Let $\mathcal{T} = \{T_j : T_j(\mathcal{S}) \to [0,1]\}$ be the family of functions, where $\mathcal{S}$ is a family of fuzzy summaries S. $\mathcal{T}$ is called j-tuple of qualitative criteria of $\mathcal{S}$.*

In this work, we consider that $j = 2$ (in Definition 3). T_1 denotes the truth of the summary as originally introduced by Zadeh [16]. For each object $y_i \in Y, i = 1,\ldots,n$, let $x_i = A_s(y_i)$ for $A_s \in A, s = 1,\ldots,m$. The formula for the truth function is as follows

$$T_1(S) = \mathcal{Q}\left(\frac{\sum_{i=1}^{n}\mathcal{P}(x_i)\star\mathcal{R}(x_i)}{\sum_{i=1}^{n}\mathcal{R}(x_i)}\right), \tag{12}$$

where S is a fuzzy linguistic summary, $\star$ is a t-norm.

Remark 1. We have the following formulas of often-used simpler forms of fuzzy summaries for '$\mathcal{Q}$ y's are $\mathcal{P}$':

$$T_1(\mathcal{S}) = \mathcal{Q}\left(\frac{1}{n}\sum_{i=1}^{n}\mathcal{P}(x_i)\right). \tag{13}$$

Note that (13) is a special case of (12) when $\mathcal{R}(x_i) = 1$ for all $x_i \in A_s(Y)$.

We also adapt the degree of support (T_2) that indicates how many objects in the dataset are covered by the particular summary, and it is defined as follows

$$T_2(\mathcal{S}) = \frac{1}{n}\sum_{i=1}^{n}\{x_i : \mathcal{P}(x_i) > 0 \wedge \mathcal{R}(x_i) > 0\}. \tag{14}$$

4 Fuzzy Linguistic Summaries as Explanations of Hidden Markov Models

Let us assume we observe a time series $\mathbf{x}=\{x_i\}, i = 1, \ldots, t$ and states $(\mathcal{O}_1, \ldots, \mathcal{O}_t)$. Our goal is two-fold: (1) find semi-continuous HMM $\lambda = (\mathbf{A}, \mathbf{B}, \pi)$ for the observed sample; (2) construct fuzzy linguistic summaries about this model and accuracy of the predictions. Let us also assume that N, M are given (Eq. (6)). Next, we run the Baum-Welch algorithm (see Section 2) to estimate the following parameters of the semi-continuous HMM λ :

- initial probabilities $\pi_i, i = 1, \ldots, N$,
- elements of the transition matrix $a_{ij}, i, j = 1, \ldots, N$,
- parameters for normal distributions $w_{jk}, \mu_k, \Sigma_k, k = 1, \ldots, M, j = 1, \ldots, N$.

Let μ_k^* and σ_k^*, $k = 1, \ldots, M$ denote the estimated parameters for the mixture of normal distribution from Eq. (6). In this work, we propose to use these estimated parameters for the construction of linguistic terms sets.

Let us remember that our goal is to build fuzzy linguistic summaries that improve the understanding of the inference process and best describe the relationships between the observations ($\mathbf{x}$) and the accuracy of the predicted states, the respective linguistic variables are denoted as A_{x_k} and A_{EMR}. For this purpose, we build fuzzy linguistic summary S (see Def. 2) of the form:

$$\mathcal{Q}\ \mathcal{R}\ \textit{objects are}\ \mathcal{P}, \tag{15}$$

where

- qualifier $\mathcal{R}$ corresponds to the values of observations $\mathbf{x}$ conditional on $k = 1, \ldots, M$ states, e.g., A_{x_k}={*below value* z_k, *around value* z_k, *above value* z_k} $z_k \in R$. Tab. 1 presents the construction of the respective fuzzy numbers;
- summarizer $\mathcal{P}$ reflects the quality of the state prediction measured with the exact match rate, e.g., A_{EMR}={*high exact match rate, low exact match rate*};
- quantifier $\mathcal{Q}$ reflecting the relative quantifiers, e.g., *almost all, majority, around half, minority, almost none.*

Table 1. Membership functions for term sets for attribute A_1 describing variable x as triangular fuzzy numbers $[a_1, a_2, a_3]$.

Linguistic term	type	a_i	b_i	c_i
below μ_k^*	z-shape		$\mu_k^* - \sigma_k^*$	μ_k^*
around μ_k^*	triangular	$\mu_k^* - \sigma_k^*$	μ_k^*	$\mu_k^* + \sigma_k^*$
above μ_k^*	s-shape	μ_k^*	$\mu_k^* + \sigma_k^*$	

As depicted in Tab. 1, the *below* terms are expressed with z-shape fuzzy numbers and are characterized by the two parameters $\mu_k^* - \sigma_k^*$ and μ_k^* ; *around* terms are expressed with triangular fuzzy numbers that are characterized by the three parameters $\mu_k^* - \sigma_k^*$, μ_k^* and $\mu_k^* + \sigma_k^*$; and the *above* terms are expressed s-shape fuzzy numbers that are characterized by the two parameters μ_k^* and $\mu_k^* + \sigma_k^*$.

Let TP denote the number of true positives, TN - the number of true negatives, FP - the number of false positives, and FN - the number of false negatives. We measure the accuracy of classification with the Exact Match Ratio (EMR) according to the following formula:

$$\text{EMR} = \frac{TP + TN}{TP + TN + FN + FP}. \tag{16}$$

Alg. 1 summarized the proposed approach for the generation of fuzzy linguistic summaries.

Algorithm 1 Generation of fuzzy linguistic summaries of hidden Markov models

Input: observations $\mathbf{x}=\{x_i\}, i = 1, \ldots, t$ and states $(\mathcal{O}_1, \ldots, \mathcal{O}_t)$, N, M
Results: fuzzy linguistic summaries $\{\mathcal{S}\}$

procedure STEP1: DEFINITIONS
  Define quantifiers Q, default value is $Q = \{most\}$
  Define attributes for performance, default is $A_{EMR} = \{$*high exact match rate, low exact match rate*$\}$
  Define quality measures, by default is the degree of truth and the degree of support T_1, T_2
end procedure
procedure STEP2: ESTIMATE PARAMETERS OF π
  $[\mu^*, \sigma^*, \mathcal{O}^*, a^*]$ = Baum-Welch algorithm ($\mathbf{x}$, $\mathcal{O}$, N, M)
  Calculate EMR (Eq. (16))
end procedure
procedure STEP3: CREATE S FOR ALL COMBINATIONS OF ATTRIBUTES
  for k in $\{1, \ldots, M\}$ **do**
    for j in $\{1, \ldots, q\}$ **do**
      for i in $\{1, \ldots, N\}$ **do**
        Construct fuzzy numbers as in Tab 1 for z_i
        Apply quantifier Q_j
        Construct S: Q_j $A(z_i)$ *objects in state* k *are* A_{EMR}
        Calculate quality measures $T_1(\mathcal{S}), T_2(\mathcal{S})$
      end for
    end for
  end for
end procedure
procedure STEP4: PRESENT OUTPUTS
  Select the most true summaries which $T_1(\mathcal{S}) = 1$
end procedure

5 Simulation Experiment

We perform a simulation experiment as an illustrative example of the performance of the proposed fuzzy linguistic summaries about HMMs described in Section 4. Let us simulate observations from (X_1, X_2, X_3) as follows

$$X_1, X_3 \sim \mathcal{N}(n = (50, 50), \mu = 0, \sigma = 1),$$
$$X_2 \sim \mathcal{N}(n = 50, \mu = 1.1, \sigma = 1).$$

The resulting time series $\mathbf{x}$ (in total 150) and corresponding $M = 2$ states are depicted in Fig. 1.

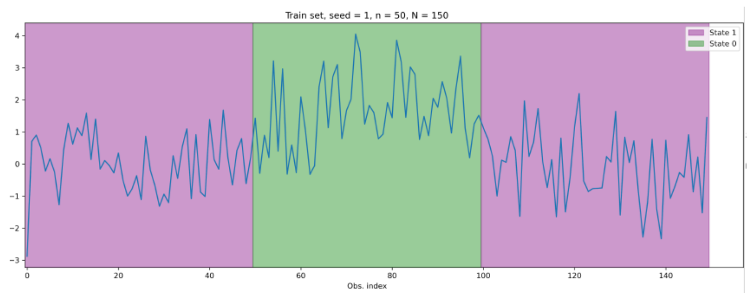

Fig. 1. Simulated data (X_1, X_2, X_3) for this illustrative experiment.

Next, we apply Baum-Welch algorithm to estimate the parameters of the semi-continuous hidden Markov model π^*. Among them are the conditional pdfs, as depicted in Figure 2.

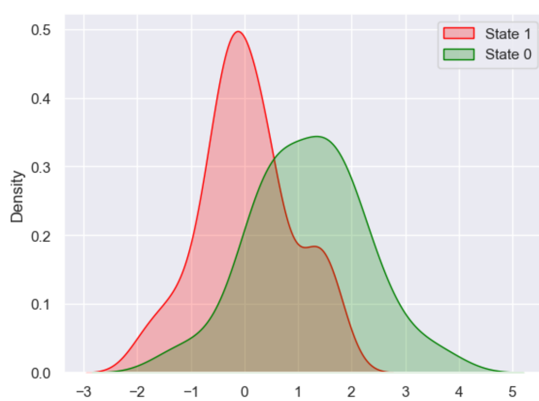

Fig. 2. Estimated pdfs for $\mathbf{x}$ in $M = 2$ states.

As observed, the estimated μ_0^* amounts to 1.1 and μ_1^* is -0.08. Next, we construct fuzzy numbers representing the linguistic terms for A_x according to Tab. 1. The resulting linguistic variables are presented in Fig. 3.

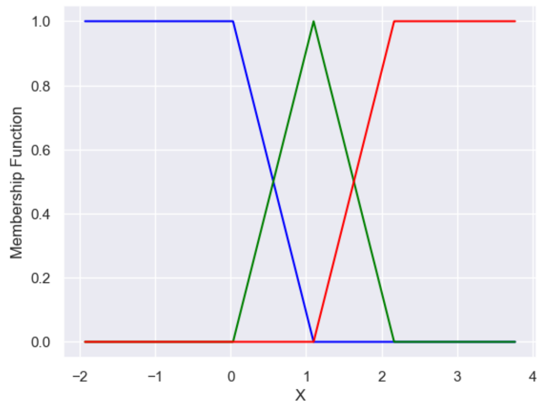

Fig. 3. Illustrative example of linguistic variables describing the values of parameter x inspired by the estimated pdfs for state $k = 0$ that is z_0=1.1.

Next, we create linguistic variables for the quantifiers and performance using A_{EMR}. In this example, we define quantifier *most* as the trapezoidal fuzzy number $[0.5, 0.8, 1, 1]$.

Finally, sample fuzzy linguistic summaries for z_0=1.1 and $T_1 = 1$ are presented in Tab. 2.

Table 2. Sample fuzzy linguistic summaries describing the relationships between the observed time series **x** and the performance of the estimated HMM measured with the exact match rate.

Fuzzy linguistic summary	T_1	T_2
For most obs. in state 0 with X lower than 1.1, exact match rate is high	1	0.44
For most obs. with X lower than 1.1, exact match rate is high	1	0.71
For most obs. in state 0 with X around 1.1 we have high exact match rate.	1	0.68
For most obs. with X around 1.1 we have high exact match rate.	1	0.53
For most obs. in state 0 with X higher than 1.1 we have high exact match rate.	1	0.46
For most obs. with X higher than 1.1 we have high exact match rate.	1	0.25

We can see that, for example, a sentence: *"For most observations in state 0 with X around 1.1 we have high exact match rate"* (with support $T_2 = 0.68$). For comparative purposes, let us now see the Shapley Additive Explanation [11] as benchmark for this simulated dataset in Fig. 4 which informs mainly about the feature importance.

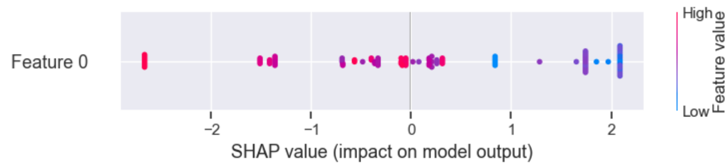

Fig. 4. Shapley plots illustrating the contribution of observations from $\mathbf{x}$ to prediction of State 0. Each point represents a classified observation and the color code represents its range of feature values.

Interestingly, in Tab. 2 we also see the sentence: *"For most observations with X around 1.1, we have high exact match rate"* the support is lower and amounts to $T_2 = 0.53$. Such insights are intuitive for the expert aware of the pdfs, however, are not possible to be easily revealed with the graphical outputs of the current XAI algorithms.

6 Conclusions and Future Work

In this work, we introduced a method enabling the construction of linguistically quantified sentences linking both the observed time series and the estimated hidden Markov model. With a simulation experiment, we showed the usefulness of linguistic summaries for explaining time series data. Specifically, we consider the semi-continuous hidden Markov models that combine a mixture of normal distributions.

The presented initial results are promising, though further research is needed to confirm all the advantages and limitations of the proposed approach. Further development of the methodology aims to explain all parameters of the hidden Markov model, particularly the transition matrix and the other characteristics of the conditional pdfs, which will allow us to express the modelling process and its results in more detail in natural language.

References

1. Alvey, B., Anderson, D., Keller, J., Buck, A.: Linguistic Explanations of Black Box Deep Learning Detectors on Simulated Aerial Drone Imagery. Sensors **23**(15), 6879 (2023)
2. Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., Herrera, F.: Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion **58**, 82–115 (2020)
3. Baum, L.: An Inequality and Association Maximization Technique in Statistical Estimation for Probabilistic Functions of Markov Processes. Inequality **3**, 1–8 (1972)

4. Gilpin, L.H., Bau, D., Yuan, B.Z., Bajwa, A., Specter, M., Kagal, L.: Explaining Explanations: An Overview of Interpretability of Machine Learning. In: 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA). pp. 80–89. IEEE, Turin, Italy (Oct 2018)
5. Huang, X.: Semi-continuous hidden Markov models for speech recognition. Phd thesis, University of Edinburgh, Edinburgh (1989)
6. Kacprzyk, J., Wilbik, A., Zadrożny, S.: Linguistic summarization of time series using a fuzzy quantifier driven aggregation. Fuzzy Sets and Systems **159**(12), 1485–1499 (2008)
7. Kacprzyk, J., Yager, R., Merigo, J.: Towards Human-Centric Aggregation via Ordered Weighted Aggregation Operators and Linguistic Data Summaries: A New Perspective on Zadeh's Inspirations. IEEE Computational Intelligence Magazine **14**(1), 16–30 (2019)
8. Kacprzyk, J., Zadrożny, S.: Linguistic database summaries and their protoforms: towards natural language based knowledge discovery tools. Information Sciences **173**(4), 281–304 (2005)
9. Kaczmarek-Majer, K., Casalino, G., Castellano, G., Dominiak, M., Hryniewicz, O., Kamińska, O., Vessio, G., Díaz-Rodríguez, N.: PLENARY: Explaining black-box models in natural language through fuzzy linguistic summaries. Information Sciences **614**, 374–399 (2022)
10. Kaczmarek-Majer, K., Hryniewicz, O., Opara, K., Radziszewska, W., Olwert, A., Owsiński, J., Zadrożny, S.: Control Charts Designed Using Model Averaging Approach for Phase Change Detection in Bipolar Disorder. In: Destercke, S., Denoeux, T., Gil, M., Grzegorzewski, P., Hryniewicz, O. (eds.) Uncertainty Modelling in Data Science, Advances in Intell. Systems and Computing, vol. 832, pp. 115–123. Springer International (2019)
11. Lundberg, S.M., Lee, S.I.: A unified Approach to Interpreting Model Predictions. In: 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA. pp. 1–10 (2017)
12. Miś, K., Kaczmarek-Majer, K., Baczyński, M.: On some properties of fuzzy linguistic summaries. In: Book of Abstracts of The Seventeenth International Conference on Fuzzy Set Theory and Applications, FSTA 2024, Liptovský Jăn, Slovakia, January 28 - February 2, 2024. pp. 74–75. University of Ostrava (2024)
13. Rabiner, L., Juang, B.: An introduction to hidden markov models. IEEE ASSP Magazine **3**(1), 4–16 (1986)
14. Yager, R.: A New Approach to the Summarization of Data. Information Sciences **28**(1), 69–86 (1982)
15. Yager, R.: Connectives and quantifiers in fuzzy sets. Fuzzy Sets and Systems **40**(1), 39–75 (1991)
16. Zadeh, L.: A computational approach to fuzzy quantifiers in natural languages. Computational linguistics pp. 149–184 (1983)