Available online at www.sciencedirect.com
ScienceDirect
Fuzzy Sets and Systems 281 (2015) 292–299
FUZZY sets and systems
www.elsevier.com/locate/fss

Does machine learning need fuzzy logic?

Eyke Hüllermeier

Department of Computer Science, University of Paderborn, Germany

Received 27 July 2015; received in revised form 1 September 2015; accepted 2 September 2015

Available online 7 September 2015

Abstract

This article is a short position paper in which the author outlines his (necessarily subjective) perception of current research in fuzzy machine learning, that is, the use of formal concepts and mathematical tools from fuzzy sets and fuzzy logic in the field of machine learning. The paper starts with a critical appraisal of previous contributions to fuzzy machine learning and ends with a suggestion of some directions for future work.

Keywords: Fuzzy sets; Fuzzy logic; Machine learning

1. Introduction

Since their inception 50 years ago, marked by Lotfi Zadeh's seminal paper [32], and rapid emergence in the following decades, fuzzy sets and fuzzy logic have found their way into numerous fields of application, such as engineering and control, operations research and optimization, databases and information retrieval, data analysis and statistics, just to name a few.

More recently, fuzzy concepts have also been used in machine learning, giving birth to the field of *fuzzy machine learning*. This development has largely been triggered by the increasing popularity of machine learning as a key methodology of artificial intelligence (AI), modern information technology and the data sciences. Moreover, it has come along with a shift from *knowedge-based* to *data-driven* fuzzy modeling, i.e., from the manual design of fuzzy systems by human experts to the automatic construction of such systems by fitting (fuzzy) models to data.

In more classical applications like information processing and expert systems, fuzzy logic is primarily used for the purpose of knowledge representation, and inference is mostly of a *deductive* nature. Machine learning, on the other hand, is mainly concerned with *inductive* inference, namely, the induction of general, idealized models from specific, empirical data. Thus, while the key importance of probability theory and statistics as mathematical foundations of machine learning is immediately understandable and indisputable, the role of fuzzy logic in this field is arguably much less obvious at first sight.

E-mail address: eyke@upb.de.

http://dx.doi.org/10.1016/j.fss.2015.09.001
0165-0114/

The goal of this paper is to outline the author's perception of current research in fuzzy machine learning, which includes the discussion of the role of fuzzy sets in machine learning. This perception is based on significant experience with both research communities, fuzzy logic and machine learning, not only as an author of research papers but also as a reviewer, conference organizer and journal editor. In spite of this, it goes without saying that the presentation will necessarily remain subjective and potentially biased.

Prior to proceeding, it should be emphasized that this is not a survey paper. In fact, references are rather sparse and sometimes deliberately omitted (especially in connection with more critical comments or negative remarks), and the fraction of self-citations is higher than usual. Moreover, the focus of this paper is more on machine learning (model induction) and less on data mining (pattern mining, exploratory data analysis, data description). Although both fields are closely connected, there are nevertheless important differences between them, and these differences are not unimportant with regard to the possible role and potential contributions of fuzzy logic—see [16] for a more detailed discussion of this point.

2. The status quo

Inside the fuzzy (logic) community, fuzzy machine learning can nowadays be seen as an established subfield. The number of publications on this topic is still not as high as for some other subfields, such as fuzzy control, but notable and continuously increasing. There is a EUSFLAT working group on Machine Learning and Data Mining[1] and an IEEE CIS Task Force on Machine Learning.[2] Moreover, special sessions on this topic are organized quite regularly on the occasion of fuzzy conferences every year, just like special issues in journals.

That being said, the connection between the fuzzy and the core machine learning community is not well established at all. On the contrary, the two communities seem to be sharply separated, with very little (if any) interaction in the form of joint meetings, research initiatives or mutual conference attendance. For example, contributions on fuzzy machine learning are almost exclusively published in fuzzy journals and conferences, whereas it is extremely difficult to find a fuzzy paper in a core machine learning conference or journal.

Related to the lack of communication between the communities, the recognition of fuzzy logic inside machine learning is still rather moderate, to put it mildly. To some extent, this might be explained by the general reservation of AI scholars against fuzzy logic, which has diminished but not fully disappeared, as well as the fact that most ML researchers, while being well trained in probability and statistics, are still quite unfamiliar with the basics of fuzzy logic. Honestly, however, reasons can also be found on the side of the fuzzy community. For some reason, many fuzzy papers admittedly fall short of the scientific standards in machine learning, which have continuously increased over the past decades. Without attempting to give an explanation, this reflects the author's personal impression based on reviewing and editorial experience. Moreover, the fuzzy ML community seems to be somewhat lagging behind in terms of timeliness. Machine learning has developed quite rapidly in the recent past, and "hot topics" are changing quickly. While ML scholars are focusing on topics such as deep learning, manifold learning, structured output prediction, sparsity and compressed sensing, constructive induction, etc., the majority of fuzzy papers is still about rule induction, a topic that matured and essentially stopped in ML research in the 1990s.

In the following sections, existing work in fuzzy machine learning and contributions often emphasized by fuzzy scholars will be discussed in some more detail, albeit more in an exemplary rather than a comprehensive way. The next section comments on model fuzzification, because this is what most papers on fuzzy machine learning are about. Section 4 addresses the aspect of interpretability, which is typically highlighted as the main advantage of fuzzy approaches, while Section 5 is devoted to the representation of uncertainty in machine learning. Some interesting directions for future work are sketched in Section 6, prior to concluding the paper in Section 7.

3. Fuzzification of models

The bulk of contributions in fuzzy machine learning deals with the fuzzy extension of standard, non-fuzzy methods: from rule induction to fuzzy rule induction [19,14,5], from decision trees to fuzzy decisions trees [20,27,26],

1 http://www.eusflat.org/research_wg_dami.php.
2 http://cis.ieee.org/emergent-technologies-tc.html.

from nearest neighbor estimation to fuzzy nearest neighbor estimation [21], from support vector machines to fuzzy support vector machines [25,1], etc. In general, this means an extension of the representation of corresponding models by means of fuzzy concepts, such as the use of fuzzy instead of crisp partitions in decision tree learning. The effect is an increased flexibility of the model class, which can indeed be useful and improve performance, especially if the original class is quite restricted. For example, fuzzy rule induction gets rid of restrictions to axis-parallel decision boundaries, which may improve classification accuracy [15], and the finer granularity of scores produced by fuzzy decision trees (compared to standard decision trees, which produce many ties) can be useful in ranking [18].

In spite of potential advantages of that kind, the "fuzzification" of conventional machine learning methods can be questioned for several reasons:

- The intellectual challenge is typically not very high, and hence the scientific contribution not very deep. In fact, a fuzzification of a model class can normally be done in a rather straightforward way, and the necessary extension of the corresponding learning algorithms is often not very difficult either.
- The large majority of "fuzzy models" eventually implements a standard function that maps crisp input to crisp output values—at least subsequent to a defuzzification step. For example, a "fuzzy classifier" normally maps instances to class labels, just like any other (non-fuzzy) classifier; in particular, it is neither dealing with fuzzy data nor does it produce fuzzy predictions. Therefore, the true benefit in comparison to conventional models is often not very obvious. As mentioned earlier, the effect of increased flexibility could be an advantage, however, the same effect could also be achieved by passing to a more flexible non-fuzzy model class (e.g., using SVMs with Gaussian instead of linear kernels); besides, as is known from learning theory, more flexibility may as well turn out to be disadvantageous (mainly due to the risk of overfitting the training data, and therefore generalizing poorly beyond this data).
- The previous concern is amplified if fuzzy extensions cause an increased computational complexity, which is often the case.
- In some cases, the link to fuzzy sets (let alone fuzzy logic) appears to be somewhat artificial, especially if membership functions are merely used as a kind of weighting function (like in fuzzy nearest neighbor or SVM methods, for example). Sometimes, the label "fuzzy logic" is even badly misused for "playing with numbers in the unit interval". Anyway, even in cases where the notion of *fuzzy sets* can be justified, most of the methods have very little to do with *fuzzy logic*, since generalized logical operators or fuzzy inference techniques are not needed (fuzzy rule and decision tree learning are notable exceptions).[3]

4. The myth of interpretability

Interpretability is one of the core arguments often put forward by fuzzy scholars in favor of fuzzy models—usually in a very uncritical way. In fact, many authors seem to take it for granted that fuzzy models or, more specifically, fuzzy rule-based models, can easily be understood and interpreted by a human user or data analyst. Many of these authors apparently equate "fuzzy" with "linguistic" and "linguistic" with "interpretable", which, of course, if far too simple. A real "proof" of interpretability would require the presentation and careful inspection of a fuzzy model learned from data [2,3], which is almost never done. At best, a discussion of that kind is replaced by the computation of certain *interpretability measures* [8], which, however, are disputable and pretend to a level of objectivity that is arguably not warranted for this criterion.

Without denying the potential usefulness of fuzzy logic in constructing interpretable models in general, the author is quite convinced that current fuzzy machine learning methods produce models that are not at all more interpretable than any other types of model, and often even less. There are several reasons for this scepticism:

- First, there is the persistent conception that logical structures such as rules can be understood by people quite easily while analytical expressions and "formulas" cannot. One may wonder, then, why logistic regression is a standard approach for predictive modeling in the medical domain, where interpretability is indeed critical,

[3] For example, one may wonder what is actually "fuzzy" about the celebrated fuzzy c-means (FCM) clustering algorithm [6].

whereas rule models are almost absent in this field. Medical scientists are fond of logistic regression because each model parameter has a clear meaning: it quantifies the influence of an individual predictor variable, in terms of direction and strength, on the overall outcome or prediction. For example, to what extent does smoking increase the probability of lung cancer?

- Even if a single rule might be understandable, a rule-based model with a certain level of accuracy will typically consist of many such rules. For grid-based methods, for example, which are often used in the fuzzy community, the number of rules grows exponentially with the dimensionality of the input space. Who is able to digest a model with hundreds of rules? On top of this, fuzzy models often allow for rule weighting and may involve complicated inference schemes for aggregating the outputs of individual rules into an overall prediction.
  Of course, the problem of complexity is shared by other, non-fuzzy methods, too. For example, decision trees are often praised as being highly interpretable, also in mainstream ML. This might indeed be true as long as trees are sufficiently small. In real applications, however, accurate trees are often large, comprising hundred of nodes. Again, interpretability is highly compromised then.
- When fuzzy sets are constructed in a data-driven way, it is not at all clear that these sets can be associated with meaningful linguistic labels—let alone labels the semantic interpretation of which will be shared among different users. In fact, one should realize that the fuzzy sets produced are strongly influenced by the data set, which is a random sample. Consequently, the partition becomes random itself, and may agree with any "true" semantics at best by chance. Besides, a data distribution will normally not coincide with meaningful "semantic clusters". For example, the statistical distribution of the age of people in a population is far from a trimodal distribution suggesting a separation into "young", "middle-aged" and "old" people. Surprisingly, these problems are almost never discussed in papers on fuzzy machine learning.
- Likewise, it is rarely discussed in which way a model is eventually presented to the user. Are the fuzzy sets specified in terms of their membership functions, in which case the user might be overloaded with mathematical details, or is it just presented in a linguistic form? As mentioned in the previous item, the latter presupposes an appropriate assignment of linguistic labels to fuzzy sets. But even then, it might be difficult to capture the meaning, unless the user is a specialist in the domain with a clear conception of the labels' semantic. For example, what is the meaning of the rule antecedent "petal length is medium" in the famous IRIS data? You cannot have the slightest clue unless being a botanist with sufficient background knowledge about iris flowers. Compared to this, a simple interval-based antecedent such as "petal length $\in [3.5, 5.1]$" is much more unambiguous and hence more understandable, even if the sharp interval boundaries might be seen as unnatural.

It is important to realize the crucial difference between *knowledge-based* and *data-driven* fuzzy modeling. The former is at the origin of fuzzy rule-based systems and closely connected to the classical expert systems paradigm: A human expert seeks to formalize her knowledge about a functional dependency (for example, a control function) using if–then rules, taking advantage of fuzzy sets as a convenient interface between a qualitative, symbolic and a quantitative, numerical level of knowledge representation. Obviously, a model thus produced is understandable, as it is very close to what the expert has in mind. Besides, models of that kind are typically small, comprising only a few input variables. Anyway, this is where the alleged interpretability of fuzzy systems seems to come from. However, it cannot simply be transferred to the case of data-driven fuzzy modeling, where the human is not at the origin of the model but changes her role from the "producer" to the "consumer" of a model.

Needless to say, this section is not meant as an in-depth discussion of interpretability in data-driven fuzzy modeling, which would require a more differentiated view. For example, interpretability may refer to different things: Do we seek to understand a model as a whole, or are we just interested in explaining a specific prediction? Likewise, a distinction is commonly made between so-called low-level and high-level interpretability [33]. Moreover, it is indisputable that interesting contributions on the topic have been made in recent years [4]. Yet, the author insists on the claim that interpretability of fuzzy models is far from self-evident, and that current fuzzy machine learning methods produce models that tend to be less rather than more interpretable than non-fuzzy models.

Finally, one may add that, in mainstream ML, the criterion of interpretability only plays a secondary role. To some extent, this is due to the difficulty of measuring this criterion in an objective way. More importantly, however, interpretability is indeed less an issue in most of the modern applications of ML—with only few notably exceptions, such as medical data analysis.

5. Uncertainty

Fuzzy sets are connected to uncertainty modeling via possibility theory, that is, via the interpretation of membership functions as possibility distributions. More generally, alternative uncertainty formalisms, typically based on non-additive measures, have been studied extensively in the fuzzy community. Since uncertainty is inherent in inductive inference and, therefore, learning from data is inseparably connected with uncertainty, this is another opportunity to contribute to machine learning.

In fact, even if the machine learning community is still very much focused on probability as the ultimate tool for uncertainty handling, one may well argue that probability alone is not enough to capture all sorts of uncertainty relevant to learning from data. For example, a distinction between *epistemic* and *aleatoric* uncertainty in supervised learning has recently been suggested in [30]. Roughly speaking, while aleatoric uncertainty refers to the inherently non-deterministic dependency between input and output variables and, therefore, is appropriately modeled in a probabilistic way, epistemic uncertainty refers to the incomplete knowledge—due to a lack of data—about the true dependency; this partial ignorance can be modeled conveniently in terms of (fuzzy) sets of candidate models.

Many papers on fuzzy machine learning claim to handle uncertainty in a proper way. Just like for interpretability, however, this claim seems to be often taken for granted and is not substantiated by convincing arguments or sound theoretical foundations. For example, it is often not clear what kind of uncertainty is captured by a fuzzy model, and what is the meaning of a membership degree. While a *non-frequentist* interpretation of membership degrees is appealing and in a sense necessary if a demarcation from probability is sought, it seems to be a major obstacle at the same time. In particular, a non-frequentist interpretation of degrees of belief or confidence seems to be more difficult to grasp and comprehend by people than a frequentist interpretation.[4] Connected to this, the empirical evaluation of a model—a point machine learning puts much emphasis on—becomes more difficult: How to check the validity of "fuzzy predictions" empirically?

An increasing number of publications is also devoted to learning from "fuzzy data", where observations are modeled in terms of fuzzy subsets of the original data space [12,13,9,31]. Obviously, this requires the extension of corresponding learning algorithms, which normally assume precise data. Unfortunately, this is again often done without clarifying the actual meaning of a fuzzy observation and the interpretation of membership functions. In particular, just like in fuzzy statistics [23,24,28,22,10], a distinction should be made between an "ontic" interpretation, considering fuzzy data as actual entities, and an "epistemic" interpretation, considering fuzzy data as an uncertain description of some true data that is not known exactly or cannot be observed precisely [11]. In fact, these two interpretations call for fundamentally different extensions of learning methods: While the ontic interpretation suggests "lifting" standard learning methods to an extended (fuzzy) data space (via the extension principle), the epistemic interpretation will normally lead to finding a model in the original data space that is to some extent consistent with the constraints imposed by the fuzzy observations [17].

By now, the idea of learning from fuzzy data is also still hampered by the limited availability of such data, which is arguably hard to produce: Except for the human expert or data analyst himself, who may "draw fuzzy sets by hand", there does not seem to be any technical device that allows for producing fuzzy data in a systematic and automated way. Sometimes, it suggested to reinterpret a collection of [0, 1]-valued data, such as the pixels of a gray-scale image, as a single fuzzy observation. While this might be possible, and perhaps even useful in some cases, a reinterpretation of that kind is nevertheless arguable, and its advantages are often not very clear. Interestingly, an alternative way of producing and using fuzzy data was suggested in [17]: Instead of assuming fuzzy data to be given right away, precise data is systematically "fuzzified" so as to modulate the influence of individual observations on the process of model induction. Thus, the notion of "fuzzy data" is used here as a tool for modeling.

6. Quo vadis?

In the author's opinion, work in fuzzy machine learning so far has too much focused on problems for which contributions are doubtful and advantages of fuzzy extensions marginal at the best, whereas not enough emphasis has

[4] In fact, one cannot deny that thinking in terms of relative frequencies is very natural, and that proportions offer the most intuitive interpretation of numbers such as probabilities.

been put on directions for which the potential appears to be higher. Three of these directions will be briefly sketched in the following.

6.1. Modeling

Prior to the actual process of model induction, the learning problem needs to be formalized and modeled appropriately; typically, this includes the specification of various data spaces, the mathematical structure of these spaces, the relationship between them, etc. Successful learning requires a suitable formalization of the problem, which is a point that is often overlooked in machine learning. Fuzzy logic has much to offer in this regard, and definitely more than what has been realized so far.

First, one should note that "fuzzy modeling" is not restricted to expressing functional dependencies (fuzzy rules, fuzzy decision trees). Instead, there is much more that can be modeled and formalized in terms of fuzzy concepts (aggregation functions, similarity relations, etc.), albeit in a more subtle way. For example, we already mentioned the idea of "modeling data" [17]. Likewise, the structure of underlying data spaces can be characterized in terms of fuzzy relations, for example by equipping them with fuzzy order relations [7,29], on which learning algorithms can then operate conveniently.

Second, modeling becomes even more an issue in settings that go beyond standard supervised learning, such as constructive induction or reinforcement learning—in the latter, for example, a suitable abstraction of the state space is of critical importance.

6.2. Non-inductive inference

Successful principles of inductive inference, such as maximum likelihood estimation, minimum description length or structural risk minimization, are deeply rooted in probability and statistics. Yet, not all inference in machine learning is of inductive nature. For example, in problems such as multi-task learning or transfer learning, the goal is to take advantage of what has been learned in one domain while learning in another domain. For example, learning to drive a truck is easier for someone who already knows how to drive a car. The corresponding process of *knowledge transfer*, i.e., of transferring knowledge from one domain to another, perhaps only partly and not necessarily one-to-one, is largely *similarity-based* or *analogical*. Obviously, formal reasoning of that kind can nicely be supported by fuzzy inference techniques.

6.3. Uncertainty

As already said, the representation and handling of uncertainty seems to be one of the topics to which the fuzzy community can contribute the most. Possibility theory and related (even more expressive) uncertainty formalisms, such as belief functions, imprecise probabilities, etc., can complement probability theory in a reasonable way, because not all types of uncertainty relevant to machine learning are probabilistic.

Convincing ML scholars of the usefulness of such formalisms might also be a bit easier than in other cases. Despite the fact that most of them are trained in probability, it seems they are not as dogmatic as, for example, pure statisticians. At least, the relevance of capturing uncertainty is completely unquestioned, and the path from probability to formalisms that can be seen as generalizations thereof appears to be somewhat easier to tread than changing the logical way of thinking and subscribing to the idea of partial truth.

7. Concluding remarks

Given that machine learning is flourishing since many years, essentially without taking much (or even any) notice of fuzzy logic, it would be unwarranted to answer the question raised in the title of this paper with a clear "yes". Indeed, there is neither a visible contribution to the foundations of machine learning (computational/algorithmic/statistical learning theory) nor a "killer application" of machine learning that has only become possible through the support by fuzzy logic.

Having said that, it cannot be denied that interesting contributions to fuzzy machine learning have already been made and, as suggested above, even more significant ones are conceivable. This, however, presupposes fuzzy scholars to focus on the right topics, to correctly appraise the role of fuzzy sets in learning from data, to go beyond straightforward fuzzy extensions of conventional ML models, and to avoid or at least scrutinize arguable claims such as interpretability.

Not less importantly, the fuzzy community needs to get out of its isolation. This problem does not seem to be limited to machine learning but similarly applies to other application domains. There is a tendency to create a closed fuzzy X community (fuzzy control, fuzzy databases, fuzzy image processing, etc.) in parallel to the core X community. It is completely legitimate to have journals and conferences specifically devoted to fuzzy sets, just like there are journals and conferences on, say, probability theory and statistics. However, the enormous success of statistics is due to the fact that it has been established as a methodological basis in many application domains: the existence of mathematical statistics is largely justified by the multitude of applied statistics.

There is no doubt that fuzzy logic has a role to play, too, not only in mathematics, computational sciences and engineering, but in science in general. Yet, in addition to the fundamental problem and mammoth task of teaching people to think in terms of "partial truth" (a true paradigm shift in the history of science), claiming this role requires applications that show the merits of fuzzy logic in a convincing way. There are many research fields willing to accept pragmatic arguments of that kind. However, researchers from these fields are not part of the fuzzy community, and typically even unaware of fuzzy logic. Therefore, we cannot expect them to adopt fuzzy logic methodology by themselves, let alone to attend fuzzy logic conferences to learn about our tools. Instead, it is our task to make them aware of fuzzy logic and convince them of its advantages—at least if we want fuzzy logic to survive next to other mathematical an AI methodologies in the long run.

References

[1] S. Abe, Support Vector Machines for Pattern Classification, 2nd edition, Springer-Verlag, London, 2010.
[2] J.M. Alonso, L. Magdalena, A conceptual framework for understanding a fuzzy system, in: J.P. Carvalho, D. Dubois, U. Kaymak, J.M.C. Sousa (Eds.), Proc. IFSA/EUSFLAT-2009, Lisbon, Portugal, 2009, pp. 119–124.
[3] J.M. Alonso, L. Magdalena, An experimental study on the interpretability of fuzzy systems, in: J.P. Carvalho, D. Dubois, U. Kaymak, J.M.C. Sousa (Eds.), Proc. IFSA/EUSFLAT-2009, Lisbon, Portugal, 2009, pp. 125–130.
[4] J.M. Alonso, L. Magdalena, G. Gonzalez-Rodriguez, Looking for a good fuzzy system interpretability index: an experimental approach, Int. J. Approx. Reason. 51 (1) (2009) 115–134.
[5] F.J. Berlanga, A.J. Rivera Rivas, M.J. del Jesús, F. Herrera GP-COACH, Genetic programming-based learning of compact and accurate fuzzy rule-based classification systems for high-dimensional problems, Inf. Sci. 180 (8) (2010) 1183–1200.
[6] J.C. Bezdek, R. Ehrlich, W. Full FCM, The fuzzy c-means clustering algorithm, Comput. Geosci. 10 (2–3) (1984) 191–203.
[7] U. Bodenhofer, F. Klawonn, Robust rank correlation coefficients on the basis of fuzzy orderings: initial steps, Mathw. Soft Comput. 15 (2008) 5–20.
[8] J. Casillas, O. Cordon, F. Herrera, L. Magdalena (Eds.), Interpretability Issues in Fuzzy Modeling, Studies in Fuzziness and Soft Computing, Springer-Verlag, Berlin, 2003.
[9] T. Cour, B. Sapp, B. Taskar, Learning from partial labels, J. Mach. Learn. Res. 12 (2011) 1501–1536.
[10] I. Couso, D. Dubois, On the variability of the concept of variance for fuzzy random variables, IEEE Trans. Fuzzy Syst. 17 (2009) 1070–1080.
[11] I. Couso, D. Dubois, Statistical reasoning with set-valued information: ontic vs. epistemic views, Int. J. Approx. Reason. 55 (7) (2014) 1502–1518.
[12] T. Denoeux, Maximum likelihood estimation from fuzzy data using the EM algorithm, Fuzzy Sets Syst. 183 (1) (2011) 72–91.
[13] T. Denoeux, Maximum likelihood estimation from uncertain data in the belief function framework, IEEE Trans. Knowl. Data Eng. 25 (1) (2013) 119–130.
[14] A. Fernandez, M.J. del Jesús, F. Herrera, Hierarchical fuzzy rule based classification systems with genetic rule selection for imbalanced data-sets, Int. J. Approx. Reason. 50 (3) (2009) 561–577.
[15] J. Hühn, E. Hüllermeier, FURIA: an algorithm for unordered fuzzy rule induction, Data Min. Knowl. Discov. 19 (2009) 293–319.
[16] E. Hüllermeier, Fuzzy machine learning and data mining, WIREs Data Min. Knowl. Discov. 1 (4) (2011) 269–283.
[17] E. Hüllermeier, Learning from imprecise and fuzzy observations: data disambiguation through generalized loss minimization, Int. J. Approx. Reason. 55 (7) (2014) 1519–1534.
[18] E. Hüllermeier, S. Vanderlooy, Why fuzzy decision trees are good rankers, IEEE Trans. Fuzzy Syst. 17 (6) (2009) 1233–1244.
[19] H. Ishibuchi, T. Yamamoto, Rule weight specification in fuzzy rule-based classification systems, IEEE Trans. Fuzzy Syst. 13 (4) (2005) 428–436.
[20] C.Z. Janikow, Fuzzy decision trees: issues and methods, IEEE Trans. Syst. Man Cybern. 28 (1) (1998) 1–14.
[21] J.M. Keller, M.R. Gray, J.A. Givens, A fuzzy k-nearest neighbor algorithm, IEEE Trans. Syst. Man Cybern. SMC-15 (4) (1985) 580–585.
[22] R. Kruse, D. Meyer, Statistics with Vague Data, D. Reidel, Dordrecht, 1987.
[23] H. Kwakernaak, Fuzzy random variables I: definitions and theorems, Inf. Sci. 15 (1978) 1–29.
[24] H. Kwakernaak, Fuzzy random variables II: algorithms and examples for the discrete case, Inf. Sci. 17 (1979) 253–278.
[25] C.F. Lin, S.D. Wang, Fuzzy support vector machines, IEEE Trans. Neural Netw. 13 (2) (2002) 464–471.

[26] C. Marsala, Data mining with ensembles of fuzzy decision trees, in: Proc. CIDM, IEEE Symposium on Computational Intelligence and Data Mining, Nashville, TN, USA, 2009, pp. 348–354.
[27] C. Olaru, L. Wehenkel, A complete fuzzy decision tree technique, Fuzzy Sets Syst. 138 (2) (2003) 221–254.
[28] M.L. Puri, D.A. Ralescu, Fuzzy random variables, J. Math. Anal. Appl. 114 (1986) 409–422.
[29] M. Dolorez Ruiz, E. Hüllermeier, A formal and empirical analysis of the fuzzy gamma rank correlation coefficient, Inf. Sci. 206 (2012) 1–17.
[30] R. Senge, S. Bösner, K. Dembczynski, J. Haasenritter, O. Hirsch, N. Donner-Banzhoff, E. Hüllermeier, Reliable classification: learning classifiers that distinguish aleatoric and epistemic uncertainty, Inf. Sci. 255 (2014) 16–29.
[31] R. Viertl, Statistical Methods for Fuzzy Data, Wiley, 2011.
[32] L.A. Zadeh, Fuzzy sets, Inf. Control 8 (3) (1965) 338–353.
[33] S.M. Zhou, J.Q. Gan, Low-level interpretability and high-level interpretability: a unified view of data-driven interpretable fuzzy system modelling, Fuzzy Sets Syst. 159 (23) (2008) 3091–3131.