Ensemble Enhanced Evidential k-NN classifier through random subspaces

Asma Trabelsi, Zied Elouedi, Eric Lefevre

Ensemble Enhanced Evidential k-NN classifier through random subspaces

Asma Trabelsi[1,2], Zied Elouedi[1], and Eric Lefevre[2]

[1] Université de Tunis, Institut Supérieur de Gestion de Tunis, LARODEC , Tunisia
trabelsyasma@gmail.com,zied.elouedi@gmx.fr
[2] Univ. Artois, EA 3926, Laboratoire de Génie Informatique et d'Automatique de l'Artois (LGI2A), Béthune, F-62400, France
eric.lefevre@univ-artois.fr

Abstract. The process of combining an ensemble of classifiers has been deemed to be an efficient way for improving the performance of several classification problems. The Random Subspace Method, that consists of training a set of classifiers on different subsets of the feature space, has been shown to be effective in increasing the accuracy of classifiers, notably the nearest neighbor one. Since, in several real world domains, data can also be suffered from several aspects of uncertainty, including incompleteness and inconsistency, an Enhanced Evidential k-Nearest Neighbor classifier has been recently introduced to deal with the uncertainty pervading both the attribute values and the classifier outputs within the belief function framework. Thus, in this paper, we are based primarily on the Enhanced Evidential k-Nearest Neighbor classifier to construct an ensemble pattern classification system. More precisely, we adopt the Random Subspace Method in our context to build ensemble classifiers with imperfect data.

Keywords: Classifier ensemble, Random Subspace Method, Enhanced Evidential k-NN, belief function theory

1 Introduction

The core purpose of an ensemble classifier is to achieve a high accuracy for a given classification problem. The process of building an ensemble learning consists firstly of generating a set of base/weak classifiers from the training data and then perform actual classification by combining the output predictions of base classifiers. To gain a better accuracy, the basic classifiers should be diverse and independent [13]. Several ensemble classifier generation methods allow to achieve diversity among the base classifiers. Bagging [3] and Boosting [18] are widely used as ensemble methods but some authors have proven that these two techniques are not guaranteed to produce fully independent individual base classifiers [5]. Both theoretical and experimental researches conducted by the machine learning community have shown that the efficient method for achieving a good diversity consists of training the base classifiers on different feature subsets

[4, 24]. This may be explained by the fact that a feature subset-based ensemble can reduce the correlation among the classifiers and also perform faster owing to the reduced size of input features [4, 8, 11]. The key problem of this kind of ensemble learning is how to yield attribute subsets with good predicting power. Several feature subsets techniques have been introduced till now where some of which are based on filter approaches [17], while others are relied on wrapper approaches [12]. Another more popular and effective tool is the Random Subspace Method (RSM) also called random subspacing [19] and has satisfactory yielded results particulary with the standard k-Nearest Neighbor classifier (k-NN)[2]. In this paper, we have to adapt the random subspace method in the real context of uncertain data. Precisely, we propose to design a classifier ensemble via random subspacing on the basis of the Enhanced Evidential k-NN (EEk-NN) classifier, which is proposed in [23], as a new technique for dealing with uncertain data represented within the belief function framework. The reminder of this paper is organized as follows: Section 2 is committed to highlighting the fundamental concepts of the belief function theory. In Section 3, we present the EEk-NN classifier that handles evidential databases. Section 4 is dedicated to describing our proposed ensemble classifier through random subspaces. Our experimentation on several synthetic databases is conducted in Section 5. Finally, the conclusion and our main future work directions are reported in Section 6.

2 Belief function theory: background

The belief function theory, also referred to as evidence theory, is widely regarded as very effective and efficient basis for representing, managing and reasoning about uncertain knowledge. This section briefly reviews some important concepts underlying this theory.

2.1 Information representation

Let $\Theta = \{\theta_1, \theta_2, \ldots, \theta_N\}$ denote the frame of discernment including a finite non empty set of N elementary hypotheses that are assumed to be exhaustive and mutually exhaustive. The power set of Θ, denoted by 2^Θ, is made up of all the subsets of Θ:

$$2^\Theta = \{\emptyset, \theta_1, \theta_2, \ldots, \theta_N, \ldots, \Theta\} \tag{1}$$

Expert's beliefs over the subsets of the frame of discernment Θ are represented by the so-called basic belief assignment (bba) denoted by m. It is carried out in the following manner:

$$\sum_{A \subseteq \Theta} m(A) = 1 \tag{2}$$

Each subset A of 2^Θ having fulfilled $m(A) > 0$ is called a focal element.

2.2 Combination operators

For certain real world problems, we are clearly confronted with information issued from several sources. Therefore, a number of combination rules has been proposed and discussed for some past time. The conjunctive rule, introduced by Smets within the Transferable Belief Model (TBM) [21], is one of the best known ones. Given two information sources S_1 and S_2 with respectively m_1 and m_2 as bbas, the conjunctive rule, denoted by $\bigcirc\!\!\!\!\cap$, was established as follows:

$$m_1 \bigcirc\!\!\!\!\cap\, m_2(A) = \sum_{B \cap C = A} m_1(B) m_2(C), \quad \forall A \subseteq \Theta. \tag{3}$$

The belief completely associated to the empty set was recognised under the name of conflictual mass. A normalized version of the conjunctive rule has been proposed by Dempster [6] to retain the basic characteristics of the belief function theory. Indeed, it allows to manage the conflict while redistributing the conflictual mass over all focal elements. The Dempster rule is then set as follows:

$$m_1 \oplus m_2(A) = \frac{1}{1-K} \sum_{B \cap C = A} m_1(B) m_2(C), \quad \forall A \subseteq \Theta \tag{4}$$

where the conflictual mass K caused by the combination of the two bbas m_1 and m_2 through the conjunctive rule, is given as follows:

$$K = \sum_{B \cap C = \emptyset} m_1(B) m_2(C) \tag{5}$$

2.3 Decision making

The pignistic probability, denoted by $BetP$, has been proven to be an effective and efficient decision-making tool for selecting the most likely hypothesis relative to a given problem [20]. It consists of transforming beliefs into probability measures as follows:

$$BetP(A) = \sum_{B \cap A = \emptyset} \frac{|A \cap B|}{|B|} m(B), \quad \forall A \in \Theta \tag{6}$$

The hypothesis H_s that has to be chosen is the one with the highest pignistic probability:

$$H_s = argmax_A BetP(A), \quad \forall A \in \Theta \tag{7}$$

2.4 Dissimilarity between bbas

In the research literature, several measures have been proposed to compute the degree of dissimilarity between two given bbas [10, 16, 22]. One of the earliest and best-known measures is the Jousselme distance. Formally, the Jousselme distance, for two given bbas m_1 and m_2, is defined by:

$$dist(m_1, m_2) = \sqrt{\frac{1}{2}(m_1 - m_2)^T D(m_1 - m_2)} \tag{8}$$

where the Jaccard similarity measure D is set to:

$$D(X, Y) = \begin{cases} 1 & \text{if } X{=}Y{=}\ \emptyset \\ \dfrac{|X \cap Y|}{|X \cup Y|} & \forall\ X{,}Y \in 2^{\Theta} \end{cases} \tag{9}$$

3 Nearest Neighbor classifiers for uncertain data

Data uncertainty is regarded as one of the main issues of several real world applications that can affect experts' decisions. Two levels of uncertainty can be distinguished in the literature: the uncertainty that occurs in the attribute values and the one pervading the class labels. The process of constructing classifiers from totally uncertain data has not received the great attention till now. Drawing inspiration from the Evidential theoretic k-NN that incorporates classifier outputs uncertainties [7, 9], we have proposed an EEk-NN classifier for handling not only the uncertainty associated with the classifier outputs but also that pervading the data, precisely the attribute values. Suppose we have to solve an M class classification problem. Let us denote by $X = \{x^i = (x^i_1, ..., x^i_n); L^i | i = 1, ..., N\}$ a collection on N n-dimensional training samples where each one is characterized by n uncertain attribute values x^i_j $(j \in \{1, \ldots, n\})$ represented within the belief function framework and a class label L^i demonstrating its membership to a specific class in $\Theta = \{\theta_1, \ldots, \theta_M\}$. Assume that $y{=}\{y_1{,}\ldots{,}y_n\}$ be a new query pattern to be classified on the basis of the training set X. The major idea underlying our proposed classifier is to compute the distance $d_{y,i}$ between the query pattern y and each instance $x^i \in X$ that corresponds to the sum of the absolute differences between the attribute values as follows:

$$d_{y,i} = \sum_{j=1}^{n} dist(x^i_j, y_j) \tag{10}$$

Particulary, we have relied on the Jousselme distance measure $dist$ (see Equation 8) for processing the uncertainty that characterizes the attribute values. It must be emphasised that $d_{y,i}$ can have values comprised within the range of from 0 to 1. A value of $d_{y,i}$ which is too small involves the situation that the instances y and x^i are described by the same class label L^i. In contrast, a high value of $d_{y,i}$ implies the situation of almost complete ignorance with regard to the class label of y. As a matter of fact, the uncertainty pervading the class label of the query pattern y can be modeled and represented within the belief function theory. Assume that the training instances are sorted in ascending order according to their distance from the test instance y, each training instance $x^i \in X$ provides an item of evidence denoted by $m^{(i)}(.|x^i)$ over Θ:

$$\begin{aligned} &m^{(i)}(\{\theta_q\}|x^i) = \alpha\Phi_q(d_{y,i}) \\ &m^{(i)}(\Theta|x^i) = 1 - \alpha\Phi_q(d_{y,i}) \\ &m^{(i)}(A|x^i) = 0, \forall A \in 2^{\Theta} \backslash \{\Theta, \theta_q\} \end{aligned} \tag{11}$$

where the distance function $d_{y,i}$ should be calculated such as in Equation 8, θ_q refers to the class label of x^i and α is a parameter satisfying $0 < \alpha < 1$. It has been proven that a value of α equal to 0.95 can lead to satisfactory or better outcomes [7]. The decreasing function Φ_q, checking $\Phi_q(0)$=1 and $lim_{d \to \infty}\Phi_q(d) = 0$, will be given as follows:

$$\Phi_q(d) = exp(-\gamma_q d^2), \tag{12}$$

where γ_q displays a positive parameter of class θ_q. It can be optimized depending on the training samples. An exact method relied on a gradient search procedure can be used for small or medium data sets, while using a linearization approach for large data [25]. The best values of γ_q, for both exact and approximated methods, can be estimated by minimizing the mean squared classification error over the whole training set X of size N.

The final bba m^y regarding the class membership of the query pattern y can be obtained by merging the bbas issued from k nearest neighbors training instances of y through the Dempster rule of combination. The final bba will be defined as follows:

$$m^y = m^{(1)}(.|x^1) \oplus m^{(2)}(.|x^2) \oplus \ldots \oplus m^{(k)}(.|x^k) \tag{13}$$

The class label concerning the test pattern y, will be made by computing the pignistic probability $BetP$ of the bba m^y as shown in Equation 6. The query pattern y is then assigned to the class label with the highest pignistic probability.

4 Ensemble Enhanced Evidential *k*-NN (Ensemble EE*k*-NN)

As already mentioned, the concept of diversity is regarded as a vital necessity for the ultimate success of ensemble classifier systems. Note however, that in this context, the RSM is a widely used technique addressed to ensure diversity between individual classifiers and has achieved satisfactory results notably for the ensembles of Evidential theoretic k-NNs [1]. Despite their relevance and success, such kind of ensemble systems cannot handle imperfect data, especially the uncertain ones. Get inspired from [1], in this paper, we propose a new ensemble system that fully benefits from the advantages of both RSM and EEk-NN. Our proposed ensemble classification system deals mainly with uncertain data where the uncertainty occurs precisely in the attribute values and is represented within the belief function framework. The suggested model is generally characterised by three main steps. Given a training data X, the first level concerns the generation of T feature subsets with size S from a uniform distribution over X. In the second level, the output label of each query pattern will be predicted through T EEk-NN classifiers that are trained with the different generated feature subsets. The final stage concerns the combination of the predictions yielded by the different classifiers. Let us remind that the output label of each individual classifier is expressed in terms of a mass function. The belief function theory has also been proven to be an efficient way for merging an ensemble of classifiers where each of which produces a belief function for each query instance. Different combination rules have been implemented within this framework and can be categorized according to the dependency between the merged sources. In this paper, we ultimately opted for the Dempster operator, which is the conventionally used rule within the belief function theory, for combining diverse classifiers.

Two substantial parameters need to be considered for our proposed framework:

- **The number of created classifiers:** A substantial key element when designing an ensemble classifiers is the number of individual classifiers used to get the final decision. There is no doubt that a huge number of classifiers may in the one hand increase the computational complexity and on the other hand decrease the comprehensibility. Several researches have been done to predefine a reasonable number of classifiers. The conclusion conducted following to the study of [15] shows that ensembles of 25 k-NN classifiers are sufficient for reducing the error rate and consequently for improving performance. For that very reason, in this paper, we set the number of combined EEk-NN classifiers to 25.

- **The size of feature subsets S:** The choice of the appropriate size of feature subsets is still being studied. Since a small subspace size can make the algorithm even faster, the chance to fall into missing informative features or also missing correlation between several features can ever be strong enough. To address that challenge, in this paper, we will randomly select the subspace size, relative to each individual EEk-NN classifier, in the range $[n/3;2n/3]$, which means that at least one-third and at most two-thirds of the original feature set will be used to train each component classifier (i.e. the subspace size S varies from one classifier to another).

5 Experimentations

This Section is devoted to studying the performance improvements of our Ensemble EEk-NN classifier in random subspaces compared with that in full feature space. Our comparative study will mainly be based on the percentage of correct classification (PCC) criterion. In what follows, we elaborate our experimentation settings (Section 5.1) and our experimentation results (Section 5.2).

5.1 Experimentation settings

Since we are dealing specifically with uncertain knowledge, we have generated several synthetic databases while injecting a degree of uncertainty P, having values comprised within the range [0,1], to some well−known real data sets obtained from the UCI machine learning repository [14]. Table 1 provides a short description of the different tested databases where #Instances, #Attributes and #Classes denote, respectively, the number of instances, the number of attributes and the number of classes. Four uncertainty levels have been considered in this paper: certain case (P=0), low uncertainty case($0 < P < 0.4$), middle uncertainty case ($0.4 \leq P < 0.7$) and high uncertainty case ($0.7 \leq P \leq 1$).

Table 1: Description of databases

Databases	#Instances	#Attributes	#Classes
Voting Records	435	16	2
Heart	267	22	2
Monks	195	23	2
Lymphography	148	18	4
Audiology	226	69	24

Let D be a given database described by N instances x^i ($i \in \{1, \ldots, N\}$) and n attributes x^i_j ($j \in \{1, \ldots, n\}$). Let Θ_j be the frame of discernment associated to the attribute j. Suppose that $|\Theta_j|$ is the cardinality of Θ_j, every attribute value $v^i_{j,t}$ relative to an instance x^i such that $v^i_{j,t} \subseteq \Theta_j$ ($t \in \{1, \ldots, |\Theta_j|\}$) will be represented through the belief function framework as follows:

$$\begin{aligned} &m^{\Theta_j}\{x^i\}(v^i_{j,t}) = 1 - P \\ &m^{\Theta_j}\{x^i\}(\Theta_j) = P \end{aligned} \tag{14}$$

5.2 Experimentation results

To assesses our model performance, we have undertaken the 10-fold cross validation strategy. This technique splits randomly the treated data into ten equal sized parts where nine part is used as a training set and the remaining as testing sets. A major key issue in our proposed approach is related to the number of neighbors that may give satisfactory results, in our current experimentation tests, we evaluate five values of the nearest neighbors k which respectively correspond to 1, 3, 5, 7 and 9. The PCC results are given from Table 2 to Table 6.

Table 2: Results for Heart database (%)

	$k=1$		$k=3$		$k=5$		$k=7$		$k=9$	
	EEk-NN	Ensemble EEk-NN	EEk-NN	Ensemble EE$k-$NN	EEk-NN	Ensemble EEk-NN	EEk-NN	Ensemble EE$k-$NN	EE$k-$NN	Ensemble EE$k-$NN
No	61.15	**67.30**	63.84	**70.38**	67.30	**68.07**	70	**70.03**	71.15	**71.23**
Low	58.46	**68.84**	64.23	**66.15**	66.92	**69.23**	68.07	**68.07**	**79.03**	78.24
Middle	60	**69.23**	63.07	**65.38**	66.15	**67.69**	**69.61**	67.30	**68.07**	67.69
High	63.84	**68.46**	63.07	**65.76**	66.36	**66.53**	70.76	**71.13**	69.61	**70.03**

Table 3: Results for Vote Records database (%)

	$k=1$		$k=3$		$k=5$		$k=7$		$k=9$	
	EEk-NN	Ensemble EEk-NN	EEk-NN	Ensemble EE$k-$NN	EEk-NN	Ensemble EEk-NN	EEk-NN	Ensemble EE$k-$NN	EE$k-$NN	Ensemble EE$k-$NN
No	92.79	**92.05**	92.32	**92.65**	**93.02**	92.32	93.72	**94.01**	**93.72**	92.81
Low	92.09	**93.14**	93.02	**93.65**	92.55	**93.24**	93.25	**94.25**	93.25	**94.78**
Middle	91.62	**92.79**	91.39	**92.56**	91.39	**93.12**	91.86	**92.94**	92.32	**94.16**
High	84.18	**87.20**	87.67	**88.60**	88.60	**89.30**	**89.30**	86.97	89.76	**91.86**

Table 4: Results for Monks database (%)

	$k=1$		$k=3$		$k=5$		$k=7$		$k=9$	
	EEk-NN	Ensemble EEk-NN	EEk-NN	Ensemble EE$k-$NN	EEk-NN	Ensemble EEk-NN	EEk-NN	Ensemble EE$k-$NN	EE$k-$NN	Ensemble EE$k-$NN
No	72	**73.13**	59.81	**60.26**	60.54	**61.68**	**70**	69.03	79.81	**80.45**
Low	69.63	**71.01**	58.18	**59.49**	63.63	**94.16**	**70.90**	70.65	76.54	**77.88**
Middle	68.9	**69.85**	63.81	**64.23**	66.72	**68.9**	71.09	**72.84**	70.72	**72.13**
High	54.90	**56.14**	53.09	**53.68**	**52.54**	52.03	52.72	**53.26**	54.18	**55.36**

Table 5: Results for Audiology database (%)

	$k=1$		$k=3$		$k=5$		$k=7$		$k=9$	
	EEk-NN	Ensemble EEk-NN	EEk-NN	Ensemble EE$k-$NN	EEk-NN	Ensemble EEk-NN	EEk-NN	Ensemble EE$k-$NN	EE$k-$NN	Ensemble EE$k-$NN
No	63.18	**64.22**	60.45	**60.67**	52.72	**53.16**	50.45	**51.26**	44.54	**45.22**
Low	52.72	**52.98**	55.45	**55.67**	53.63	**53.87**	47.27	**47.56**	45.9	**46.81**
Middle	52.72	**53.24**	**48.18**	47.84	**44.54**	44.22	41.13	**42.76**	40.45	**41.68**
High	**15.45**	14.49	23.18	**24.01**	21.36	**22.45**	22.27	**23.46**	18.18	**18.96**

Table 6: Results for Lymphography database (%)

	$k = 1$		$k = 3$		$k = 5$		$k = 7$		$k = 9$	
	EEk-NN	Ensemble EEk-NN	EEk-NN	Ensemble EE$k-$NN	EEk-NN	Ensemble EEk-NN	EEk-NN	Ensemble EE$k-$NN	EE$k-$NN	Ensemble EE$k-$NN
No	84.28	**84.51**	85	**85.07**	62.42	**63.45**	85.71	**86.25**	85	**85.42**
Low	80	**81.12**	85.71	**86.13**	**83.57**	82.56	**86.42**	81.17	85.71	**86.96**
Middle	82.14	**82.42**	**84.28**	83.96	86.42	**87.22**	84.28	**85.14**	82.14	**83.27**
High	58.57	**58.63**	61.42	**62.19**	59.28	**61.02**	58.57	**59.13**	65.71	**66.48**

According to the results given from Table 2 to Table 6, we can deduce that ensembles of the EEk-NN classifier through random subspacing has led to interesting results compared to the individual EEk-NN classifiers that are learnt with the full feature space. In fact, the PCC yielded by an ensemble of classifiers is generally better than that yielded by an individual classifier for the most of cases. For instance, let us consider k equals 5, the PCC results yielded by the ensemble system on the Heart database with No, Low, Middle and High uncertainties are respectively equal to 67.30%, 66.92%, 66.15% and 66.36%. However, there are equal to 68.07%, 69.23%, 67.69% and 66.53% when using an individual system. This small difference may be explained by the existence of irrelevant and redundant features as a consequence of the random method.

6 Conclusion

In this paper, we have proposed an ensemble EEk-NN classifier through random subspaces with the aim of increasing the classification performance for a given classification problem. For assessing the performance of our proposed approach, we have carried out a comparative study between the ensemble EEk-NN classifier in random subspaces and that in full feature space when relied on the PCC assessment criterion. Although the RSM method can unfortunately increase the risk that irrelevant and redundant features may be part of the selected subsets, numerical results have shown that ensemble EEk-NN classifiers have contributed to somewhat more favorable PCC results for the different mentioned databases. To promote better and more effective classification performance, in our future studies and research projects, we look forward to solutions allowing to produce the best possible feature subsets.

References

[1] H. Altınçay. Ensembling evidential k-nearest neighbor classifiers through multi-modal perturbation. *Applied Soft Computing*, 7(3):1072–1083, 2007.

[2] S. D. Bay. Combining nearest neighbor classifiers through multiple feature subsets. In *15th International Conference on Machine Learning*, volume 98, pages 37–45, 1998.

[3] L. Breiman. Bagging predictors. *Machine learning*, 24(2):123–140, 1996.

[4] R. Bryll, R. Gutierrez-Osuna, and F. Quek. Attribute bagging: improving accuracy of classifier ensembles by using random feature subsets. *Pattern recognition*, 36(6):1291–1302, 2003.

[5] S. B. Cho and H.-H. Won. Cancer classification using ensemble of neural networks with multiple significant gene subsets. *Applied Intelligence*, 26(3):243–250, 2007.

[6] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *The annals of mathematical statistics*, pages 325–339, 1967.

[7] T. Denoeux. A k-nearest neighbor classification rule based on Dempster-Shafer theory. *IEEE transactions on systems, man, and cybernetics*, 25(5):804–813, 1995.

[8] S. Günter and H. Bunke. Feature selection algorithms for the generation of multiple classifier systems and their application to handwritten word recognition. *Pattern recognition letters*, 25(11):1323–1336, 2004.

[9] L. Jiao, T. Denœux, and Q. Pan. Evidential editing k-nearest neighbor classifier. In *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, pages 461–471. Springer, 2015.

[10] A. Jousselme, D. Grenier, and E. Bossé. A new distance between two bodies of evidence. *Information fusion*, 2(2):91–101, 2001.

[11] Y. Kim. Toward a successful crm: variable selection, sampling, and ensemble. *Decision Support Systems*, 41(2):542–553, 2006.

[12] R. Kohavi and G. H. John. Wrappers for feature subset selection. *Artificial intelligence*, 97(1-2):273–324, 1997.

[13] L. Kuncheva, M. Skurichina, and R. P. Duin. An experimental study on diversity for bagging and boosting with linear classifiers. *Information fusion*, 3(4):245–258, 2002.

[14] P. Murphy and D. Aha. UCI repository databases. *http://www.ics.uci.edu/mlear*, 1996.

[15] D. Opitz and R. Maclin. Popular ensemble methods: An empirical study. *Journal of Artificial Intelligence Research*, 11:169–198, 1999.

[16] B. Ristic and P. Smets. The TBM global distance measure for the association of uncertain combat id declarations. *Information Fusion*, 7(3):276–284, 2006.

[17] N. Sánchez-Maroño, A. Alonso-Betanzos, and M. Tombilla-Sanromán. Filter methods for feature selection–a comparative study. In *International Conference on Intelligent Data Engineering and Automated Learning*, pages 178–187. Springer, 2007.

[18] R. E. Schapire. The boosting approach to machine learning: An overview. In *Nonlinear estimation and classification*, pages 149–171. Springer, 2003.

[19] M. Skurichina and R. P. Duin. Bagging, boosting and the random subspace method for linear classifiers. *Pattern Analysis & Applications*, 5(2):121–135, 2002.

[20] P. Smets. Decision making in the TBM: the necessity of the pignistic transformation. *International Journal of Approximate Reasoning*, 38(2):133–147, 2005.

[21] P. Smets and R. Kennes. The Transferable Belief Model. *Artificial intelligence*, 66(2):191–234, 1994.

[22] B. Tessem. Approximations for efficient computation in the theory of evidence. *Artificial Intelligence*, 61(2):315–329, 1993.

[23] A. Trabelsi, Z. Elouedi, and E. Lefevre. A novel k-nn approach for data with uncertain attribute values (to appear). In *30th International Conference on Industrial, Engineering and other Applications of Applied Intelligent Systems*. Springer, 2017.

[24] K. Tumer and J. Ghosh. Classifier combining: analytical results and implications. In *Proceedings of the National Conference on Artificial Intelligence*, pages 126–132, 1996.

[25] L. M. Zouhal and T. Denoeux. An evidence-theoretic k-nn rule with parameter optimization. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 28(2):263–271, 1998.