HAL open science

Evaluation of Naive Evidential Classifier (NEC): Application to semolina milling value

Yosra Mazigh, Boutheina Ben Yaghlane, Sébastien Destercke

To cite this version:

Yosra Mazigh, Boutheina Ben Yaghlane, Sébastien Destercke. Evaluation of Naive Evidential Classifier (NEC): Application to semolina milling value. Scalable Uncertainty Management (SUM 2012), Sep 2012, Marburg, Germany. pp.619-624. hal-00745590

HAL Id: hal-00745590

https://hal.science/hal-00745590

Submitted on 25 Oct 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NAIVE EVIDENTIAL CLASSIFIER (NEC): APPLICATION TO SEMOLINA MILLING VALUE

Yosra Mazigh

LARODEC Laboratory, Institut Supérieur de Gestion de Tunis

Boutheina Ben Yaghlane

LARODEC Laboratory, Institut des Hautes Etudes Commerciales de Carthage

Sébastien Destercke

CNRS UMR 7253 Heuristique et diagnostique des systèmes complexes

We introduce the notion of naive evidential classifier. This classifier, which has a structure mirroring the naive Bayes classifier, is based on the Transferable Belief Model and use mass assignments as its uncertainty model. This new method is based achieves more robust inferences, mainly by explicitly modeling imprecision when data are in little amount or are imprecise. After introducing the model and its inference process based on Smet's generalized Bayes theorem (GBT), we specify some possible methods to learn its parameters, based on the Imprecise Dirichlet Model (IDM) or on predictive belief functions. Some experimental results on an agronomic application are then given and evaluated.

1. Introduction

When modeling and processing uncertainty, computing on multivariate spaces is an important issue. If $X_1, \ldots, X_N$ is a set of variables assuming their values over some finite spaces $\mathcal{X}_1, \ldots, \mathcal{X}_N$, defining directly a joint uncertainty model over the Cartesian product $\mathcal{X}_1 \times \ldots \times \mathcal{X}_N$ and making some inferences with this model is often impossible in practice. The use of graphical models based on network architecture can solve this problem by decomposing the joint uncertainty model into several pieces of conditional models. This decomposition is possible thanks to conditional independence property. Note that outside their computational tractability, another attractive feature of such models is their readability for non-experts, thanks to their graphical aspects.

The aim of this paper is first to introduce the notion of the Naive Ev-

idential Classifier (NEC) as the counterpart of the Naive Bayes (NB) in evidential theory. It uses the directed evidential network structure (DEVN) proposed by Ben Yaghlane in[1] and preforms inference by using the modified binary join tree (MBJT) algorithm, which uses the disjunctive rule of combination (DRC) and the generalized Bayesian theorem (GBT), both proposed by Smets in.[6]

In Section 2, we briefly present the Basics of NEC. Section 3 then provides some details about the practical instanciation and use of the NEC structure. Finally, Section 4 presents some preliminary experiments on an agronomical problem. Due to lack of spaces, only the essential elements are provided, and the reader is referred to references for details.

2. Naive Evidential Classifier (NEC)

NEC the TBM counterpart of the Naive Bayes (NB) classifier. Thereby, it is a graphical model having two parts: a qualitative and a quantitative one. Recall that the aim of a classifier is learn a mapping from input values $X_1, \ldots, X_N \in \mathcal{X}_1 \times \ldots \times \mathcal{X}_N$ to an output class $C \in \mathcal{C}$ from available (training) data, in order to predict the classes of new instances.

2.1. *Graphical structure*

NEC maintain the same graphical presentation as NB. In the figure 1, the root C is class to predict and the leafs from X_1 to X_n present features.

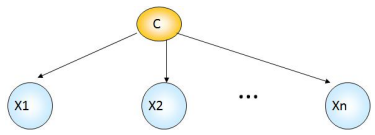

Fig. 1. A generalized presentation of the naive evidential classifier NEC

2.2. *Quantitative part of NEC*

Each edge represents a conditional relation between the two nodes it connects. To each edge will be associated conditional mass distribution*, while

*Recall that a mass distribution $m : 2^{|X|} \rightarrow [0,1]$ on $\mathcal{X}$ is such that $m(\emptyset) = 0$, $\sum_{E \subseteq \mathcal{X}} m(E) = 1$ and induces a belief and a plausibility measure such that $bel(A) = \sum_{E \subseteq A} m(A)$ and $pl(A) = 1 - bel(A^c)$.

to each node will be associated a prior mass as follows:

- A prior mass distribution $m(C)$ in the root (class) node.
- Both a prior mass distribution $m(X)$ and a conditional mass distribution $m^X[C](x_i)$ in each leaf node associated to an edge.

Section 3 explains how these masses can be learnt. The propagation of beliefs for NEC, ensured by a Modified Binary Joint Tree Structure, is illustrated in Figure 2 for nodes C and X_2. We refer to[1] for details.

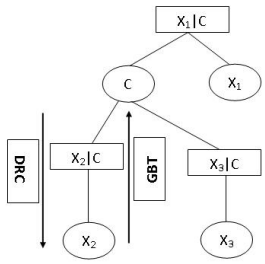

Fig. 2. The propagation process in the MBJT using the GBT and the DRC

3. Learning and decision

3.1. *Learning*

Mass distributions in the model can be elicited from experts or constructed from observed data (the usual case in classification). We propose two methods to infer such distributions: the Imprecise Dirichlet model (IDM)[2] and Denoeux's multinomial model,[4] obtaining respectively the NEC1 and NEC2 paremeterization.

The imprecision of mass distributions obtained by each methods is governed by hyperparameters, respectively the positive real number $s \in \mathbb{R}^+$ for the IDM and the confidence level $\alpha \in [0, 1]$ for Denoeux's model. The higher these parameters, the higher the imprecision of mass distributions.

3.2. *Classification and evaluation*

As imprecision is an interesting feature of evidence theory, we propose an imprecise classification based on pairwise comparisons.[5] Using the mass on

$\mathcal{C}$ obtained for an instance $x_1, \ldots, x_N$ by propagation on the NEC, a class c_1 is told to dominate c_2, denoted by $c_1 \succeq c_2$ if the belief of c_1 is larger than the belief of c_2 ($bel(c_1) \geq bel(c_2)$) and the plausibility of c_1 is larger than the plausibility of c_2 ($pl(c_1) \geq pl(c_2)$) for all the distributions. The retained set Ω_m of possible classes is then the one of non-dominated classes, i.e. $\Omega_m = \{c \in \mathcal{C} |\ \nexists c', c' \succ c\}$. Evaluations are then made through the use of set accuracy and of discounted accuracy (see[3]). Roughly speaking, if $\hat{c}$ is the true class, set-accuracy counts 1 each time $\hat{c} \in \Omega_m$, while the discounted accuracy counts $1/|\Omega_m|$ with $|\Omega_m|$ the cardinality of Ω. Both counts 0 when $\hat{c} \notin \Omega_m$, and classical accuracy is recovered when $|\Omega_m| = 1$

4. Experiments On Durum Wheat

In this section we present preliminary results on an agronomic application of NEC consisting in the prediction of the semolina milling value of the durum wheat. Being able to predict this value from easy to measure parameters would be very valuable both for farmers and industrials, the former because it would help in grain selection, the latter to quickly assess wheat quality.

The database contains 256 samples issued from IATE experimental mill, where the output class was the semolina value (discretized in four classes specified by experts) and the input parameters were the Hectolitre Weight (HLW), The Thousand Kernel Weight(TKW) and Vitreousness. NEC1 and NEC2 are both evaluated according to set and discounted accuracy measures, and a ten-fold cross validation method was used.

Figures 3 and 4 show the variations of set- and discounted- accuracy as a function of the parameters v (IDM) and α (Denoeux's model). A

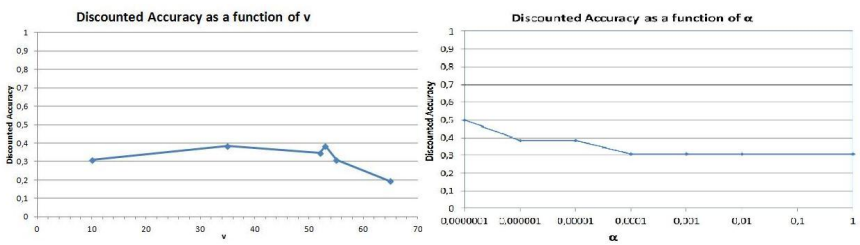

Fig. 3. Discounted Accuracy as a Function of v and α

first remark is that NEC1 tends to have worse performances than NEC2, an observation that can be explained by the fact that the IDM is more

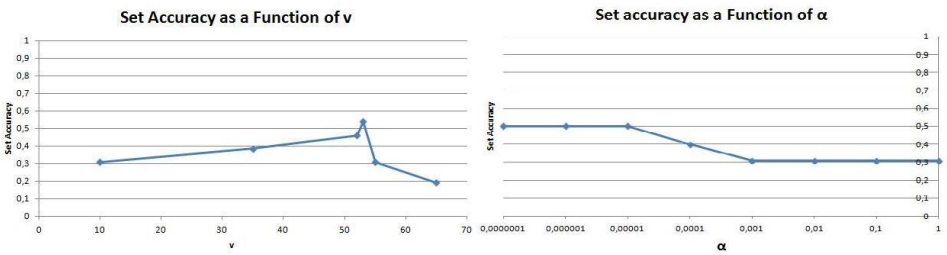

Fig. 4. Set Accuracy as a Function of v and α

"simple" than Denoeux's model. Hence NEC 1 is easier to compute, while NEC2 gives better predictions. Another noticeable fact is that set-accuracy does not necessarily decrease as the model gets more imprecise (Figure 4 left). This is due to the non-motonicity of the used decision rule (i.e., a more imprecise m does not mean a bigger Ω_m), and this could be solved by choosing another (more imprecise) rule such as interval-dominance.[5]

Figure 5 displays the confusion matrix for both classifiers (where the precise classification was chosen as the most plausible, i.e., $c = \arg\max_{c \in \mathcal{C}} pl(\{c\})$). Again, both classifiers appear to be at odds, and it is difficult to say which is better. However, as is supported by the relatively low level of well-classed items, this difficulty to differentiate probably also comes from the poor explanatory power of input variables, and more refined modelling (planned in the future) as well as tests on other data sets would be needed to make further conclusions.

Finally, to choose the best configuration, we propose to retain the best one according to discounted accuracy (as it reflects balance between imprecision and prediction quality), that is to say the NEC1 classifier with $v = 53$.

5. Conclusion

In this paper, we have introduced the idea of Naive Evidential Classifier as well as tools to instantiate and evaluate it. A preliminary application has been achieved on an agronomical problem. Considering the quality of available data, first results are encouraging, however much work remains to be done:

- concerning the model itself, it would be desirable to realx the independence assumption anduse an augmented tree model;

Prédites / Vraies		[63;70]	[70;72]	[72;74]	[74;78,3]
		1	2	3	4
[63;70]	1	13	43	13	0
[70;72]	2	11	61	14	0
[72;74]	3	12	31	27	0
[74;78,3]	0	7	19	9	0

(a)

Prédites / Vraies		[63;70]	[70;72]	[72;74]	[74;78,3]
		1	2	3	4
[63;70]	1	36	24	9	0
[70;72]	2	21	65	13	0
[72;74]	3	9	36	25	0
[74;78,3]	0	6	10	19	0

(b)

Fig. 5. Confusion Matrix: (a)given by NEC1 with $v = 53$; and (b)given by NEC2 with $\alpha = 0.0000001$

- concerning the application, a refined statistical analysis of the data may allow to extract more relevant information or identify subgroups of interest (e.g., differentiating big and small grains or varieties of wheat);
- concerning the general evaluation of the model, it remains to apply it to usual benchmarks and to confront it to other classical classifiers.

References

1. B. Ben Yaghlane and K. Mellouli. Inference in directed evidential networks based on the transferable belief model. *International Journal of Approximate Reasoning*, 48:399–418, 2008.
2. J.-M. Bernard. An introduction to the imprecise dirichlet model for multinomial data. *International Journal of Approximate Reasoning*, 39(2-3):123–150, 2005.
3. G. Corani and M. Zaffalon. Jncc2: The java implementation of naive credal classifier 2. *Journal of Machine Learning Research*, 9:2695–2698, 2008.
4. T. Denoeux. Constructing belief functions from sample data using multinomial confidence regions. *International Journal of Approximate Reasoning*, 42(3):228–252, 2006.
5. S. Destercke. *A Decision Rule for Imprecise Probabilities Based on Pair-Wise Comparison of Expectation Bounds*, pages 189–197. Springer-Verlag Berlin, 2010.
6. P. Smets. Belief Functions: the Disjunctive Rule of Combination and the Generalized Bayesian Theorem. *International Journal of Approximate Reasoning*, 9:1–35, 1993.