A Bagging method for Cost-sensitive Imprecise Classification*

Serafín Moral-García SERAMORAL@DECSAI.UGR.ES
Joaquín Abellán JABELLAN@DECSAI.UGR.ES
Department of Computer Science and Artificial Intelligence, University of Granada, Spain

Classifiers tend to predict a single value of the class variable when classifying an instance. However, sometimes, it may be more appropriate that classifiers predict a set of class values because the available information is not sufficient for pointing out a single class value. This is known as *Imprecise Classification* [4]. When an imprecise classifier is employed, a set of values of the class variable might be obtained, composed of those class values that are not defeated by another one according to a criterion called the *dominance criterion*.

Abellán and Masegosa [1] proposed an adaptation of a decision tree model based on imprecise probabilities and uncertainty measures for Imprecise Classification. Such an adaptation is called *Imprecise Credal Decision Tree* (ICDT). Afterwards, the first ensemble method for Imprecise Classification was developed [3]. It is a Bagging scheme that uses ICDT as the base classifier. To combine multiple imprecise predictions, it employs a technique that consists of predicting as non-dominated only the class values predicted as dominated by the minimum number of classifiers. In this way, it is aimed that the ensemble is as informative as possible even though it implies a higher risk of incorrect predictions.

Precise and imprecise classifiers minimize the number of misclassifications. This would be optimal if all classification errors were equally important. Nevertheless, in practical applications, different classification errors often yield different costs. For this reason, classifiers that consider the misclassification costs, called *cost-sensitive classifiers*, have been developed.

Few cost-sensitive imprecise classifiers have been proposed so far. ICDT was adapted for cost-sensitive scenarios by Abellán and Masegosa [1]. That adaptation was improved by Moral-García et al. [2]. Despite the fact that ensemble schemes have obtained better performance than individual classifiers, no ensemble method has been proposed so far for cost-sensitive Imprecise Classification. This might be because it is not trivial how to combine multiple imprecise predictions, and the matter is even more complicated when misclassification costs are different.

This poster will describe ongoing research for a Bagging method for cost-sensitive Imprecise Classification that uses the cost-sensitive Imprecise Credal Decision Tree that has obtained the best performance as the base learner. The key point is how to combine the predictions made by the multiple individual cost-sensitive imprecise classifiers. For this purpose, we aim to propose a technique that, for each class value, computes a score based on the minimum number of classifiers that predict such a value as dominated and the misclassification cost associated with that class value. Hence, we aim that the ensemble scheme is as informative as possible while we consider the error costs.

We aim to carry out an experimental analysis with many datasets and the main evaluation metrics for cost-sensitive Imprecise Classification to highlight that our proposed Bagging scheme significantly outperforms a single cost-sensitive Imprecise Credal Decision Tree.

References

[1] Joaquín Abellán and Andrés R. Masegosa. Imprecise classification with credal decision trees. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 20(5):763–787, 2012. doi:10.1142/S0218488512500353.

[2] Serafín Moral-García, Joaquín Abellán, Tahani Coolen-Maturi, and Frank P.A. Coolen. A cost-sensitive imprecise credal decision tree based on nonparametric predictive inference. *Applied Soft Computing*, 123:108916, 2022. doi:10.1016/j.asoc.2022.108916.

[3] Serafín Moral-García, Carlos J. Mantas, Javier G. Castellano, María D. Benítez, and Joaquín Abellán. Bagging of credal decision trees for imprecise classification. *Expert Systems with Applications*, 141:112944, 2020. doi:10.1016/j.eswa.2019.112944.

[4] Marco Zaffalon. The naive credal classifier. *Journal of Statistical Planning and Inference*, 105(1):5 – 21, 2002. doi:10.1016/S0378-3758(01)00201-4.

Acknowledgments This work has been supported by UGR-FEDER funds under Project A-TIC-344-UGR20, by the “FEDER/Junta de Andalucía-Consejería de Transformación Económica, Industria, Conocimiento y Universidades" under Project P20-00159, and by he Spanish Ministry of Education and Science under project PID2019-106758GB-C31.