



Evaluation in learning from label proportions: an approximation to the precision-recall curve*

*Note: The full contents of this paper have been published in the volume *Lecture Notes in Artificial Intelligence 11160* (LNAI 11160)

Jerónimo Hernández-González
Intelligent Systems Group
University of the Basque Country UPV/EHU
Donostia, Spain
jeronimo.hernandez@ehu.eus

Abstract—In the last decade, the learning from label proportions problem has attracted the attention of the machine learning community. Many learning methodologies have been proposed, although the evaluation with real label proportions data has hardly been explored. This paper proposes an adaptation of the area under the precision-recall curve metric to the problem of learning from label proportions. The actual performance is bounded by minimum and maximum approximations. Additionally, an approximate estimation which takes advantage of low-uncertain bags is proposed. The benefits of this proposal are illustrated by means of an empirical study.

Index Terms—Learning from label proportions, Weakly supervised classification, Evaluation, Precision-recall curve