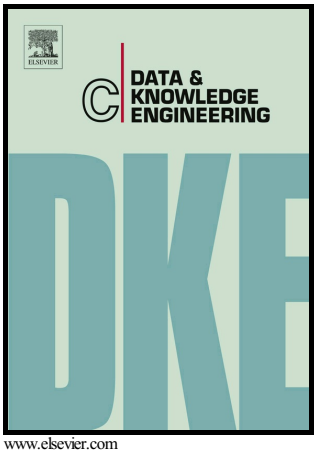

An Information-Theoretic Filter Approach for Value Weighted Classification Learning in Naive Bayes

Chang-Hwan Lee

PII: S0169-023X(16)30127-6
DOI: http://dx.doi.org/10.1016/j.datak.2017.11.002
Reference: DATAK1623

To appear in: *Data & Knowledge Engineering*

Received date: 2 August 2016
Revised date: 1 November 2017
Accepted date: 11 November 2017

Cite this article as: Chang-Hwan Lee, An Information-Theoretic Filter Approacl for Value Weighted Classification Learning in Naive Bayes, *Data & Knowledg Engineering,* http://dx.doi.org/10.1016/j.datak.2017.11.002

An Information-Theoretic Filter Approach for Value Weighted Classification Learning in Naive Bayes

Chang-Hwan Lee

Department of Information and Communications
Dongguk University, Seoul, Korea

Abstract

Assigning weights in features has been an important topic in some classification learning algorithms. In this paper, we propose a new paradigm of assigning weights in classification learning, called value weighting method. While the current weighting methods assign a weight to each feature, we assign a different weight to the values of each feature. The performance of naive Bayes learning with value weighting method is compared with that of some other traditional methods for a number of datasets. The experimental results show that the value weighting method could improve the performance of naive Bayes significantly.

Keywords: Feature Weighting, Feature Selection, Naive Bayes, Kullback-Leibler

1. Introduction

In some classifiers, the algorithms operate under the implicit assumption that all features are of equal value as far as the classification problem is concerned. However, when irrelevant and noisy features influence the learning task to the same degree as highly relevant features, the accuracy of the model is likely to deteriorate. Since the assumption that all features are equally important hardly holds true in real world application, there have been some attempts to relax this assumption in classification. Zheng and Webb [1] provide a comprehensive overview of work in this area. The first approach for relaxing this assumption is to combine feature subset selection with classification learning. It is to combine a learning method with a preprocessing step that eliminates redundant features from the data. Feature selection methods usually adopt a heuristic search in the space of feature subsets. Since the number of distinct feature subsets grows exponentially, it is not reasonable to do an exhaustive search to find optimal feature subsets. In the literature, it is known that the predictive accuracy of naive Bayes can be improved by removing redundant or highly correlated

Email address: chlee@dgu.ac.kr (Chang-Hwan Lee)