

Efficient Semi-Supervised Feature Selection: Constraint, Relevance, and Redundancy

Abstract:

This paper describes a three-level framework for semi-supervised feature selection. Most feature selection methods mainly focus on finding relevant features for optimizing high-dimensional **data**. In this paper, we show that the relevance requires two important procedures to provide an efficient feature selection in the semi-supervised context. The first one concerns the selection of pairwise constraints that can be extracted from the labeled part of **data**. The second procedure aims to reduce the redundancy that could be detected in the selected relevant features. For the relevance, we develop a filter approach based on a constrained Laplacian score. Finally, experimental results are provided to show the efficiency of our proposal in comparison with several representative methods.