Mining Semantically Consistent Patterns for Cross-View Data

ABSTRACT:

In some real world applications, like information retrieval and data classification, we often confront with the situation that the same semantic concept can be expressed using different views with similar information. Thus, how to obtain a certain Semantically Consistent Patterns (SCP) for cross-view data, which embeds the complementary information from different views, is of great importance for those applications. However, the heterogeneity among cross-view representations brings a significant challenge on mining the SCP. In this paper, we propose a general framework to discover the SCP for cross-view data. Specifically, aiming at building a feature-isomorphic space among different views, a novel Isomorphic Relevant Redundant Transformation (IRRT) is first proposed. The IRRT linearly maps multiple heterogeneous low-level feature spaces to a high-dimensional redundant feature-isomorphic one, which we name as mid-level space. Thus, much more complementary information from different views can be captured. Furthermore, to mine the semantic consistency among the isomorphic representations in the mid-level space, we propose a new Correlation-based Joint Feature Learning (CJFL) model to extract a unique high-level semantic subspace shared across the feature-isomorphicdata. Consequently, the SCP for cross-view data can be obtained. Comprehensive experiments on three datasets demonstrate the advantages of our framework in classification and retrieval.