

**CHENNAI – PONDICHERRY**

**Semi-Supervised Learning Through Label Propagation on Geodesics**

**Abstract**

Graph-based semi-supervised learning (SSL) has attracted great attention over the past decade. However, there are still several open problems in this paper, including: 1) how to construct an effective graph over data with complex distribution and 2) how to define and effectively use pair-wise similarity for robust label propagation. In this paper, we utilize a simple and effective graph construction method to construct the graph over data lying on multiple data manifolds. The method can guarantee the connectiveness between pair-wise data points. Then, the global pair-wise data similarity is naturally characterized by geodesic distance-based joint probability, where the geodesic distance is approximated by the graph distance. The new data similarity is much more effective than previous Euclidean distance-based similarities. To apply data structure for robust label propagation, Kullback-Leibler divergence is utilized to measure the inconsistency between the input pair-wise similarity and the output similarity. In order to further consider intraclass and interclass variances, a novel regularization term on sample-wise margins is introduced to the objective function. This enables the proposed method fully utilizes the input data structure and the label information for classification. An efficient optimization method and the convergence analysis have been proposed for our problem. Besides, out-of-sample extension is discussed and addressed. Comparisons with the state-of-the-art SSL methods on image classification tasks have been presented to show the effectiveness of the proposed method.