Discrete Nonnegative Spectral Clustering

ABSTRACT

- Spectral clustering has been playing a vital role in various research areas. Most traditional spectral clustering algorithms comprise two independent stages which may cause unpredictable deviation of resultant cluster labels from genuine ones, thereby leading to severe information loss and performance degradation.

- In this work, we study how to achieve discrete clustering as well as reliably generalize to unseen data.

- We propose a novel spectral clustering scheme which deeply explores cluster label properties, including discreteness, nonnegativity and discrimination, as well as learns robust out-of-sample prediction functions.

CONT..

- We explicitly enforce a discrete transformation on the intermediate continuous labels, which leads to a tractable optimization problem with a discrete solutionwe preserve the natural nonnegative characteristic of the clustering labels to enhance the interpretability of the results..

- Moreover, to further compensate the unreliability of the learned clustering labels, we integrate an adaptive robust module with $\ell 2,p$ loss to learn prediction function for grouping unseen data.

- We also show that the out-of-sample component can inject discriminative knowledge into the earning of cluster labels under certain conditions. Extensive experiments conducted on various data sets have demonstrated the superiority of our proposal as compared to several existing clustering approaches

EXISTING SYSTEM

- It optimizing the spectral clustering models will lead to a np-hard problem due to the discrete constraint on the clustering labels.

- To achieve a feasible approximate solution ,most spectral clustering algorithms follow a common practical paradigmit first relaxes discret constraint to allow the eclustering label matrix to be continuous-value and perform seigen value decomposition.

- Then we can discretize the clustering label matrix by employing certain independent technique, such as k-means .Furthermore, to enable clustering new unseen data, one may learn an additional prediction function in an independent stage.

- The out-of sample problem is addressed by introducing a regression learning module, and discriminative information is injected into the construction of the similarity matrix to improve clustering performance.

DISADVANTAGE

- High risk of severe deviation of approximate solution from the genuine discrete clustering labels

- Information loss among separate independent stages.
- Continuous label generation, label discretization and prediction function learning.

- Unreliability of the predicted cluster labels leading to poorprediction functions.

PROPOSE SYSTEM

- We introduce an efficient algorithm for optimization and analyze its convergence.
- Though the two-stage strategy provides a feasible solution, it may unpredictably deviate from the genuine discrete clustering labels.
- To avoid this situation ,we intend to devise a unified spectral clustering model to directly generate the discrete clustering label matrix.
- An efficientand effective algorithm which iteratively solves an alternative optimization problem and guarantees the obtained solution is the optimal solution to the original problem .

ADVANTAGES

- Spectral clustering has been extensively used in real-world applications, such as age/video segmentation
- Spectral clustering family lies in the exploration of the intrinsic data structures which are fully employed for predicting clustering labels by exploiting the different similarity graphs of data points
- DNC can achieve better performance than DSC

SOFTWARE REQUIREMENT

- **HARDWARE REQUIREMENT:**
- System : Pentium IV 2.4 GHz.
- Hard Disk : 40 GB.
- Floppy Drive : 1.44 Mb.
- Monitor : 15 VGA Colour.
- Mouse : Sony.
- Ram : 512 Mb.

CONT..

- **SOFTWARE REQUIREMENT:**
- Operating system : Windows XP.
- Coding Language : ASP. Net with C#
- Data Base : SQL Server 2005.

CONCULSION

- In this work, we coped with the problem existing in most traditional spectral clustering algorithms, e.G., Relaxing discrete constraints to continuous one, which consists of two independent stages (e.G., First learning continuous labels and then rounding the learned labels into discrete ones).

- In order to reduce information loss and performance degradation, we proposed a unified spectral clustering approach to directly learn discrete clustering labels and robust out-of-sample prediction functions. To be more specific, our proposed approach can explicitly rotate continuous labels to discrete ones.

- Meanwhile, we also deliberately preserv enonnegative property of cluster label matrix to get closer to the genuine partition of data. To the end of handling the noisy clustering labels, we integrated an adaptive robust module to learn prediction function for unseen data.

- Extensive experiments on six data sets demonstrated the promising performance of our proposal as compared to existing clustering approaches.