Abstract

The area of multi label text classification is getting more attention of researchers because of its role in the field of information retrieval, text mining, web mining etc. Supervised methods from machine learning are mainly used for its realization. But as it needs labeled data for classification all the time, semi supervised methods are now a day getting popular in the MLTC domain. The goal of Semi supervised learning is to reduce the classification errors using readily available unlabeled data in conjunction with available labeled data. This paper mainly provides survey and analysis of various semi supervised methods used in multi label text classification task; This overview concludes that consideration of semantic aspects of input document datasets, their representation in conjunction with smoothness and manifold assumptions in semi supervised learning may give more relevant classification results.

References

- G. Tsoumakas, G. Kalliris, and I. Vlahavas, "Effective and efficient multilabel classification in domains with large number of labels," Proc. Of the ECML/PKDD 2008 workshop on Mining Multidimensional Data (MMD'08); 08(2008) 30-44.

Index Terms

Computer Science Information Sciences
Keywords
Semi Supervised Learning  Multi Label Text Classification  Smoothness And Manifold Assumptions