Abstract

In manipulating data such as in supervised or unsupervised learning, we need to extract new features from the original features for the purpose of reducing the dimension of feature space and achieving better performance. In this paper, we investigate a novel schema for unsupervised feature extraction for classification problems. We based our method on clustering to achieve feature extraction. A new similarity measure based on trend analysis is devised to identify redundant information in the data. Clustering is then performed on the feature space. Once groups of similar features are formed, linear transformation is realized to extract a new set of features. The simulation results on classification problems for experimental data sets from UCI machine learning repository and face recognition problem show that the proposed method is effective in almost cases when compared to conventional unsupervised methods like PCA and ICA.
A New Unsupervised Clustering-based Feature Extraction Method


Index Terms

Computer Science Pattern Recognition
Keywords
Unsupervised feature extraction  similarity measure  clustering  face recognition and classification problems