M V Siva Prasad, J Data Mining In Genomics & Proteomics 2015, 6:4
http://dx.doi.org/10.4172/2153-0602.C1.003

OMICS International
conferenceseries.com

2nd International Conference on
Big Data Analysis and Data Mining
November 30-December 01, 2015 San Antonio, USA

Feature selection using an effective dimensionality reduction technique

M V Siva Prasad
Anurag Engineering College, India

Processing applications with a large number of dimensions has been a challenge to the KDD (Knowledge Discovery and Data mining) community. An effective dimensionality reduction technique is an essential pre-processing method to remove noisy features. The proposed combined method for feature selection, where a filter based on correlation is applied on whole features set to find the relevant ones, and then, on these features a wrapper is applied in order to find the best features subset for a specified predictor. The high dimensionality of data can cause data overload, and if there are a lot of features, it is possible that the number of cases in data set to be insufficient for data mining operations. The solution for these problems is the reduction of data dimensions. The size of a data set is determined both by the number of cases and by the number of features considered for each case. In order to reduce number of cases one can use sampling or filtering. Feature reduction may be achieved either by feature composition or by feature selection. These methods should produce fewer features, so the algorithms can learn faster. Sometimes, even the accuracy of built models could be improved. Methods used for feature selection, can be classified as: filters, which are open loop methods, and wrappers, which are closed loop methods. As a possibility to reduce the number of feature considered by data mining algorithms, in order to make them more efficient, the combined method which uses a combination filter wrapper. The method used a correlation based filter on the whole set of features, then on relevant subset of features it used a wrapper which uses a decision tree classifier for prediction. The combined method achieved clearly superior performances for execution time, when we have used for feature selection the combined approach and backward selection as search strategy for wrapper.

Biography

M V Siva Prasad was awarded PhD degree from Nagarjuna University, Guntur. He received MTech [SE] from VTU, Belgaum and BE [CSE] from Gulbarga University. He is presently working as Professor/Principal in Anurag Engineering College (AEC), Ananthagiri(V), Kodad(M), Nalgonda(Dt.), Telangana, India. He has published more than 50 articles in various national/international journals. He has attended 2 international conferences organized ny IACSIT and got best paper awards.

magantisivaprasad@gmail.com

Notes: