Feature selection using Yu's similarity measure and fuzzy entropy measures

Author(s)

Cesar Iyakaremye ; Pasi Luukka ; David Koloseni

Abstract

In classification problems feature selection has an important role for several reasons. It can reduce computational cost by simplifying the model. Also when the model is taken for practical use fewer inputs are needed which means in practice, that fewer measurements from new samples are needed. Removing insignificant features from the data set makes the model more transparent and more comprehensible. In this way the model can be used to provide better explanation to the medical diagnosis, which is an important requirement in medical applications. Feature selection process can also reduce noise, this way enhancing the classification accuracy. In this article feature selection method based similarity measure using Yu's similarity with fuzzy entropy measures is introduced and it is tested together with the similarity classifier. Model was tested with dermatology data set. When comparing the results to previous works the results compare quite well. Mean classification accuracy with dermatology data set was 98.83% and it was achieved using 33 features instead of 34 original features. Results can be considered quite good.