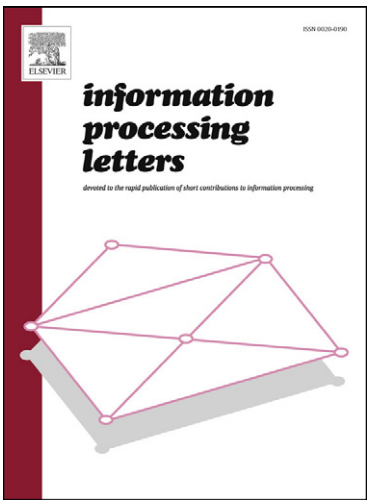
Accepted Manuscript



Efficient feature selection based on correlation measure between continuous and discrete features

Sheng-yi Jiang, Lian-xi Wang

PII: S0020-0190(15)00127-1
DOI: http://dx.doi.org/10.1016/j.ipl.2015.07.005
Reference: IPL 5307

To appear in: *Information Processing Letters*

Received date: 17 May 2013
Revised date: 6 July 2015
Accepted date: 11 July 2015

Please cite this article in press as: S.-y. Jiang, L.-x. Wang, Efficient feature selection based on correlation measure between continuous and discrete features, *Inf. Process. Lett.* (2015), http://dx.doi.org/10.1016/j.ipl.2015.07.005

Highlights

- In Section 3 of the current paper, we propose a new method that can measure the correlation (similarity) between a continuous feature and a discrete feature. To the best of our knowledge, fewer researchers attempt to measure the correlation between a continuous and a discrete feature in the literature.
- In Section 4, we elaborate on a theoretical framework to formally analyze some properties of correlation (similarity) measure between a continuous and a discrete feature. And we provide a detailed description of the proposed algorithm.
- The proposed approach *ECMBF* selects feature subset by removing irrelevant features according to the correlation between features and the class feature, and eliminating redundant features among the relevant ones.
- The proposed algorithm can efficiently and effectively achieve the high degree of dimensionality reduction and decrease classification error rates with the selected features.