

## Editorial

---

Dear Colleague:

Welcome to volume 10(6) of *Intelligent Data Analysis* – An international Journal.

Following is a brief overview of this issue that consists of 6 articles. The first three articles are about data preprocessing and understanding the contents of data, whether it is from the noise point of view or understanding the data characteristics. Of the other three papers in this issue, one is about prediction and classification, one on data mining and web maintenance and the last paper on prediction and emotional learning.

The first article by Van Hulse and Khoshgoftaar introduces a novel technique for identifying itemset observations that contain noise. The core of their approach is based on assigning a “noise factor” that would indicate a relative likelihood that an observation may contain class noise. The paper contains results of their analysis using several data sets and a bench marking of their approach with two well-known techniques for class noise identification, which are classification filter and ensemble filter. Sigut *et al.*, in the second article, propose an approach to determine whether or not a set of data comes from a univariate normal distribution. The proposed approach is based on combining neural networks to boost the power of ordinary statistics methods used in hypothesis testing. Their results show that the output of a properly trained neural nets can approximate probabilities associated with normality or non-normality with a high degree of precision. And the last article on data preprocessing is a review article by Agyemang *et al.*, where numeric and symbolic outlier mining is surveyed. Their survey discusses practical applications of outlier mining and provides a taxonomy for categorizing related mining techniques. They also provide a comprehensive review of a number of techniques discussing their advantages and disadvantages along with a discussion of some of the current research in this area.

Kant and Khan, in the fourth article of this issue, discuss classification and prediction where the focus of study is to use machine learning methods to predict pseudo-random binary sequences. The approach, which is independent of parameters and domain knowledge, suggests converting a theoretical prediction problem into a classification problem. Their article includes reports of a number of experiments where the classificatory prediction resulted in the evolution of the next bit prediction model. Web site maintainability using data mining techniques is the topic of the next paper by Burn-Thornton *et al.*, The authors show how the website pages or files, which are normally based on HTML code or as an XSL file, can be improved through analyzing certain attributes. The result of this work is also a set of style sheets which could be referred to as web page or site design in an enterprise from which the data is analyzed. The last paper of this issue by Babaie *et al.* is about prediction based on emotional learning. They propose a physiologically motivated prediction approach that is a trade off between precision and predictive performance. The particular prediction problem in this research is solar conditions and the results of their evaluation includes comparison with ground based forecasts where they demonstrate that powerful alert systems can be deigned using this approach.

And finally, with this issue of the IDA journal we celebrate the completion of our first 10 years of successful publications. We would like to greatly acknowledge all the support that we have received from the IDA community, in particular readers and authors who have made the success of this journal possible. We look forward to receiving your ideas and contributions for years to come.

With our best wishes,

Dr. A. Famili  
*Editor-in-Chief*