Transactions on Machine Learning
and Data Mining
Vol. 8, No.2 (2015) 59-60

ISSN: 1865-96781
ISBN: 978-3-942952-39-2

ibai Publishing
www.ibai-publishing.org

Editorial

Prof. Giorgio Giacinto
Universitå di Cagliari, Cagliari Italy

The human mind is continuously immersed in a flow of processing the signals captured by the five senses. The senses provide the information about the environment around us, the brain analyse the signals, and combine them in many creative ways by leveraging on previous experience. When one of the senses does not work for an illness, an injury or any other reason, then the brain makes an effort to rearrange its connections with the aim of compensating the missing signal [1].

When humans started building devices to increase the sensing abilities, than it become evident that accuracy can be attained by employing multiple sensors, effective filters, sophisticated reasoning algorithms and a large computational capability. The use of multiple sensors is brought from the human experience, as taking action by relying on the information captured by just one of the senses can be misleading.

But if humans can train themselves to capture more details on incomplete information in order to minimize misleading effects, this is quite challenging when we turn to machine data processing. Except for very simple tasks, automatic information processing requires the fusion of data from different sources. When dealing with physical sensors, this requirement translates into the need of deploying multiple types of sensors in different locations, and when we deal with information processing, this requirement can be translated into the need of acquiring data from multiple different sources. The underlying assumption is that each source conveys part of the desired information with some noise, and only the fusion of information allows gathering the multiple aspects of the desired information, while filtering out the noise. Noise suppression is possible as soon as the noise in the signal of one source is unique to that source, and, consequently, is not supported by any other source.

The literature on machine learning and data mining have tackled this issue a number of decades ago, by proposing several techniques and algorithm for fusing information from different sources [2-3]. Multiple theoretical foundations have been developed to provide solid ground in a number of industrial applications where accuracy, precision, and reduction of false positives are strict design requirements. To date, researchers and practitioners can benefit from the vast body of knowledge developed in the past years, thus providing a vast source of inspiration for novel application domains, as well as for more advanced theoretical results to extend the scope of the available techniques.

The three papers in this issue come from quite different application domains, but they share the common goal of solving the task at hand by fusing information from different sources.

The first paper proposes an ensemble method for chemical entity recognition [4]. Different information sources can be queried to obtain the desired information, and different attempts have been made to construct corpora for handling chemical-related information based on different corpus-construction guidelines. Unfortunately, each tool developed for a particular guideline might fall short when used to extract other chemical-related entities. The authors propose an original processing architecture to filter and combine information from different tools.

How to achieve accurate physical measurements in the open air is the topic addressed in the second paper in this issue [5]. The mathematical and statistical modelling that support the precision of measurements in laboratory environment need to be strengthened when used in open-air conditions, so that the concepts of traceability, uncertainty evaluation and calibration can also be applied outside of the lab. The author shows that model selection and Bayesian averaging are the two basic ingredients that help addressing this issue.

Finally, the third paper illustrates the architecture of a software tool developed to help financial analysts and investors to make sound decisions by analysing and correlating large archives of news articles [6]. The final goal of the tool is to provide a comprehensive view of events from multiple sources, thus avoiding or, at least, minimizing, bad economic consequences due to decisions based on partial information.

Prof. Giorgio Giacinto
Universitå di Cagliari, Cagliari Italy

References

1. Karns C.M., Dow M.V., Neville H.J., Altered Cross-Modal Processing in the Primary Auditory Cortex of Congenitally Deaf Adults: A Visual-Somatosensory fMRI Study with a Double-Flash Illusion, The Journal of Neuroscience, 32(28), 2012, 9626-9638
2. Khaleghi B., Khamis A., Karray F.O., Razavi S.N., Multisensor data fusion: A review of the state-of-the-art, Information Fusion, 14(1), 2013, 28-44
3. Kuncheva, L. I. Ensemble Methods, in Combining Pattern Classifiers: Methods and Algorithms, Second Edition, John Wiley & Sons, 2014.
4. Thaer M.Dieb and Masaharu Yoshioka, Extraction of Chemical and Drug Named Entities by Ensemble Learning Using Chemical NER Tools Based on Different Extraction Guidelines, Transactions on Machine learning and Data mining, Ibai-publishing, 2015, 8(2), 61-80
5. Alistair B. Forbes. Traceable Measurements using Sensor Networks, Transactions on Machine learning and Data mining, Ibai-publishing, 2015, 8(2), 81-104
6. Haiyi Zhang, Design and Considerations of a Searching Software, Transactions on Machine learning and Data mining, Ibai-publishing, 2015, 8(2), 105-104