



Editorial for the special issue: Knowledge-based systems and data science

Hamido Fujita¹ · Ali Selamat²

Published online: 7 February 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Intelligence displayed by machines has evolved from theoretical research to vital process re-improvements this very decade, where more and more industrial progression relied on the capability of machines perceiving and learning from surrounding environments, assisting in succession of given tasks. The advanced progressing of computational intelligence has promoted design and implementation of multifaceted solutions for complex industrial problems that eventually benefits changing applications in multi-disciplinarily fields. This special issue highlights the presentation of recent theoretical, operational and real-world applications on multiple industries benefitted through the implementation of artificial intelligence techniques in a vast array of domains including network security, hardware and software, data structures, ranking predictions, pattern recognition, medical tools, ubiquitous computing, education and else. The papers in this special issue have been selected from the extended versions of the 29th International Conference on Industrial, Engineering & Other Applications of Applied Intelligent Systems (IEA/AIE 2016, Morioka, Japan) and the submitted papers through a public call. All papers have been peer-reviewed according to the journal standards.

The accepted papers in this special issue are classified into three groups.

The first group focuses on the area of security, regressive methods and algorithm design. The paper titled “**Vector Based Genetic Algorithm (VBGA) to optimize predictive analysis in network security**” proposes a new Intrusion Detection System (IDS) for network security, making use of a Vector-Based Genetic Algorithm (VBGA) inspired by

evolutionary approaches. The novelty in the algorithm is to represent chromosomes as vectors and training data as matrices. This approach allows multiple pathways to calculate fitness function out of which one particular methodology is used and tested. The paper titled “**Anomaly Detection using Piecewise Aggregate Approximation in the Amplitude**” proposes a representation method called the Piecewise Aggregate Approximation in the Amplitude Domain (AD-PAA). For discovering anomalies, a sequence is partitioned into subsequences by a sliding window firstly. Then in the AD-PAA method, a subsequence is divided into equal-sized subsections according to the amplitude domain. With mean values of subsections computed, the amplitude oscillation of a subsequence is embodied effectively. The paper titled “**Supervised Ranking Framework for Relationship Prediction in Heterogeneous Information Networks**” proposes a novel frame work named Supervised Ranking frame- work (S-Rank) is proposed to tackle existing approaches. To avoid the class imbalance problem, the S-Rank framework treated the relationship prediction problem as a ranking task and divide it into three phases: Supervised PageRank strategy (SPR), utilization of Meta Path-based Ranking method (MPR) and integration of two ranking scores combining all the Attr-info, Global-info and Local-info together. Experiments on DBLP data demonstrate that the proposed S-Rank framework can effectively take advantage of all the three kinds of information for relationship prediction over HINs and outperforms other well-known baseline approaches. The paper titled “**Challenging State-of-the-Art Move Ordering with Adaptive Data Structures**” presents a comparison to two well-known, acclaimed strategies, which operate on a similar philosophy to the History-ADS, namely, the History Heuristic, and the Killer Moves technique. In this work, a mechanism by which these established move ordering strategies can be approximated, or directly implemented, in terms of ADSs. In a wide range of two-player and multi-player games, at various points in the game’s progression, the History-ADS performs at least as well as presented strategies (in the article), and, in fact, outperforms them in the majority of cases.

✉ Hamido Fujita
HFujita-799@acm.org
Ali Selamat
aselamat@utm.my

¹ Iwate Prefecture University, Takizawa, Iwate, Japan

² Universiti Teknologi Malaysia, Iskandar Puteri, Johor, Malaysia

The paper titled “**ETARM: An Efficient Top-K Association Rule Mining Algorithm**” presents a novel algorithm named ETARM (Efficient Top-k Association Rule Miner) to efficiently find the complete set of top-k association rules. The proposed algorithm integrates two novel candidate pruning properties to more effectively reduce the search space. These properties are applied during the candidate selection process to identify items that should not be used to expand a rule based on its confidence, to reduce the number of candidates. An extensive experimental evaluation on six standard benchmark datasets show that the proposed approach outperforms the state-of-the-art TopKRules algorithm both in terms of runtime and memory usage. The paper titled “**SpringBoard: game-agnostic tool for scenario editing with meta-programming support**” presents the development of SpringBoard, an open source scenario editor for games using the SpringRTS engine. Extending it to support game and level requirements is achieved with multi-level meta-programming, while still providing a system that is integrated with the GUI editor and therefore intuitive to use.

The second group of articles in this special issue demonstrates machine learning and analytical solutions. The paper titled “**Successes and challenges in developing a hybrid approach to sentiment analysis**” covers some success and learning experiences attained during the developing of a hybrid approach to Sentiment Analysis (SA) based on a Sentiment Lexicon, Semantic Rules, Negation Handling, Ambiguity Management and Linguistic Variables. The proposed hybrid method is presented and applied to two selected datasets: Movie Review and Sentiment Twitter datasets. The achieved results are compared against those obtained when Naïve Bayes (NB) and Maximum Entropy (ME) supervised machine learning classification methods are used for the same datasets. The proposed hybrid system attained higher accuracy and precision scores than NB and ME, which shows its superiority when applied to the SA problem at the sentence level. The paper titled “**Online Signature Verification by Spectrogram Analysis**” proposes a novel online signature analysis methodology for touchscreens that starts with signing an interface consisting of a signature silhouette. The frequency spectrum along the signing process is stealthily extracted and spectrograms are created by short-time Fourier transforms. Since the spectrograms are found as RGB images, providing valuable information about frequency vs time, grid histograms are formed by quantization for the real signature sample. Given the discrimination purposes, a fuzzified surface is designed for computing closeness of grid histograms. The paper titled “**A View-invariant Gait Recognition Algorithm based on a Joint-Direct Linear Discriminant Analysis**” proposes a view-invariant gait recognition algorithm, which builds a unique view invariant model taking advantages of

the dimensionality reduction provided by the Direct Linear Discriminant Analysis (DLDA). Proposed scheme is able to reduce the under-sampling problem (USP) that appears usually when the number of training samples is much smaller than the dimension of the feature space. In the paper titled “**Hybrid sentiment classification on twitter aspect-based sentiment analysis**”, a new hybrid approach for twitter aspect-based sentiment classification is proposed by embedding the feature selection method. The comparison of classification accuracy using the feature selection method Principal Component Analysis (PCA), Latent Semantic Analysis (LSA), and Random Projection (RP) have been presented in this paper. Furthermore, the evaluation with different classification algorithms also demonstrates that our new hybrid approach produced meaningful results with our own Hate Crime Twitter Sentiment (HCTS) dataset and also Stanford Twitter Sentiment (STS) and Sanders Twitter Corpus (STC) benchmark dataset which represent different domain. The paper titled “**A Fuzzy Classification Model for Myocardial Infarction Risk Assessment**” presented the development of a new clinical decision support system (CDSS) to diagnose Coronary Artery Diseases (CAD). In this paper, an automatic approach to design CDSS for CAD assessment is proposed. The proposed diagnosis model is based on Random Forest algorithm, C5.0 decision tree algorithm and Fuzzy modeling. The experimental results underscore very promising classification accuracy of 90.50% while optimizing training time using UCI (the University of California at Irvine) heart diseases datasets compared to the previously reported results. The paper titled “**Agent Systems Verification and Validation: Systematic Literature Review**” presented a study that surveys existing techniques for checking agent properties and detecting faults during design, development, and runtime phases of agent system life-cycle are presented with the aim to identify implementation gap and to derive research direction. The paper “**Some Bibliometric Procedures for Analyzing and Evaluating Research Fields**” has presented new results on bibliometric analytics that can provide a good tool to scientific research predictions.

The third group in this special issue presented works done in industrial applications. The paper titled “**Convolutional Neural Network Acceleration with Hardware/Software Co-Design**” presents a design methodology for accelerating CNNs using Hardware/Software Co-design techniques, in order to balance performance and exibility, particularly for resource-constrained systems. The methodology is applied to a gender recognition case study, using an ARM processor and FPGA fabric to create an embedded system that can process facial images in real-time. The paper titled “**Design and implementation of a web-based fuzzy expert system for diagnosing depressive disorder**” proposes a design a web-based expert system for diagnosing depression using

the fuzzy Delphi method by estimating the weights and importance of depression symptoms. Fuzzy Logic is adopted to calculate the level of depression. The proposed system appears helpful for everyone, from ordinary persons to specialists in medical environments. It can also be useful to train psychology students in the area of diagnostic reasoning. In the paper titled “**Overlapping community detection based on discrete biogeography optimization**”, an overlapping community detection algorithm based on biogeography optimization is proposed. Experiments on benchmark test data, including two synthetic networks and four real-world networks, show that the proposed algorithm can achieve results with better accuracy and stability than the compared evolutionary algorithms. The paper titled “**Mining Constrained Inter-Sequence Patterns: A Novel Approach to Cope with Item Constraints**” proposes two novel algorithms for it. First is the ISP-IC (Inter-Sequence Pattern with Item Constraint mining) algorithm based on a theorem that quickly determines whether an inter-sequence pattern satisfies the constraints. Then, we propose a way to improve the strategy of ISP-IC, which is then applied to the iISP-IC algorithm to enhance the performance of the process. Finally, piISP-IC, a parallel version of iISP-IC, will be presented. Experimental results show that piISP-IC algorithm outperforms the post-processing of the state-of-the-art method for mining inter-sequence patterns (EISP-Miner), ISP-IC, and iISP-IC algorithms in most of the cases. The paper titled “**A multi-level approach to ubiquitous modeling and solving constraints in combinatorial optimization problems in production and distribution**” presents a multi-level approach to the modeling and solving of combinatorial optimization problems. It is versatile and effective owing to the use of multi-level pre-solving and multiple paradigms, such as constraint programming, logic programming, mathematical programming, and fuzzy logic, for their complementary strengths. The capability of this framework and its advantage over mathematical programming alone or over hybrid frameworks is shown in the illustrative example,

in which combinatorial optimization is used as a benchmark to prove the effectiveness of the proposed approach. The paper titled “**A novel three-band orthogonal wavelet filter bank method for an automated identification of alcoholic EEG signals**” presented an automated system for the classification of alcoholic and normal EEG signals using a recently designed duration-bandwidth product (DBP) optimized three-band orthogonal wavelet filter bank (TBOWFB) and log-energy (LE). The results have achieved the classification accuracy (CA) of 97.08%, with ten-fold cross validation strategy. The proposed model presents a promising performance, and therefore it can be used in a practical setup to assist the medical professional in the diagnosis of alcoholism using EEG signals automatically. The paper titled “**Automated detection of retinal health using PHOG and SURF features extracted from fundus images**” proposed a novel approach to develop an automated retinal health screening system in this work. This paper discusses a retinal screening system to automatically differentiate normal image from abnormal (AMD, DR, and glaucoma) fundus images. The presented results achieved 96.21% accuracy, 95.00% sensitivity, and 97.42% specificity with ten-fold cross-validation strategy using k-nearest neighbor (kNN) classifier. This novel algorithm has high potential in the diagnosis of normal eye during the mass eye screening session or in polyclinics quickly and reliably. The paper “**The Promotion Strategy of Supply Chain Flexibility Based on Deep Belief Network**” has provided an intelligent flexible model by considering the supply side and demand side in uncertain environment using belief network.

The editors hope that this issue will provide valuable knowledge, ideas, and state-of-the-art research propagation in applied intelligence specifically in reshaping industrial advancement through constant improvements of informative intelligence. With much appreciation to the contributors of this special issue, the editors hope that the hard work and efforts will be made useful for a conclusive betterment of computational industrial solutions.