



Special issue on Latin–American computational intelligence

Pablo A. Estevez^{1,4} · Daniel Sbarbaro² · Millaray Curilem³

Received: 22 February 2023 / Accepted: 23 February 2023 / Published online: 6 March 2023
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2023

The aim of this Topical Collection is to present the latest advances in Computational Intelligence either performed in Latin America, or with the important participation of Latin–American researchers. Topics included neural and learning systems, evolutionary and swarm intelligence, and fuzzy and stochastic modelling. Originally, the idea was to publish extended papers from the best contributions to the IEEE Latin-American Conference on Computational Intelligence (IEEE LACCI 2021), held in Temuco, Chile, in November 2021, but later it was decided to open the Topical Collection to new submissions related to the conference topics. This Topical Collection received 33 submissions and accepted 11 papers. Out of these 11 papers, seven are extended versions of papers presented at IEEE LACCI 2021, and four are new submissions. The geographical distribution associated with the first authors of the papers published in this Topical Collection is the following: Brazil 5, Chile 3, Ecuador 1, Mexico 1, and Uruguay 1. Regarding gender, in two out of eleven papers (~ 18%), the first author is a woman, and in five out of eleven papers (~ 45%), a woman is an author/co-author.

The first group of four papers deals with new general-purpose machine learning models, which are described as follows:

Rojas-Morales & Riff propose focusing on the tuning process of the evolutionary calibrator EVOCA by opposite scoring. A novel opposite scoring strategy is introduced that learns from parameter configurations that produce a

positive and a negative effect in the target algorithm and yields a set of configurations used to initialize tuners. Three different algorithms are tuned: an ant colony optimization algorithm for solving the multidimensional knapsack problem, a genetic algorithm for solving landscapes that follow the NK model (N components and degree K), and a particle swarm optimization algorithm for solving continuous optimization problems.

Carino-Escobar et al. develop a novel tree ensemble that allows the automatic growth of the number of trees based on the structural diversity of the first two levels of trees' nodes. The algorithm is called feature-ranked self-growing forest. The algorithm presents several advantages over other tree ensembles, among them eliminating the need of setting the number of trees within the ensemble. The performance of the proposed algorithm is tested with 30 classification and 30 regression datasets and compared with random forests.

Ferreira et al. present an evolutionary algorithm to construct multi-output symbolic regression models. This is done by extending the single-output Kaizen programming (KP) to a multi-output KP through an island model. Each island works on a different output of the multi-output symbolic regression problem using a different fitness function, while a periodic exchange of migrants between islands makes them cooperate. The algorithm is evaluated on several benchmark functions and in the context of real-world chemical process systems, using data derived from analytical solutions.

Pilaluisa et al. propose a new approach to retrieve and further integrate tabular datasets (collections of rows and columns) using union and join operations. Both processes are carried out using a similarity measure based on contextual word embeddings, which allows finding semantically similar tables and overcoming the recall problem of lexical approaches based on string similarity. Column headings and cell values are used as contextual information, and their impact on each task is evaluated. BERT and its fine-tuned version WikiTables showed more stability and better performance than RoBERTa throughout all the experiments.

✉ Pablo A. Estevez
pestevez@cec.uchile.cl

¹ Department of Electrical Engineering, University of Chile, Santiago, Chile

² Department of Electrical Engineering, Universidad de Concepción, Concepción, Chile

³ Department of Electrical Engineering, Universidad de La Frontera, Temuco, Chile

⁴ Millennium Institute of Intelligent Healthcare Engineering, Santiago, Chile

A second group of seven papers deals with applications to a focused topic such as music genre classification, automatic fire detection, detecting deepfake images or videos, classifying mass spectrometry data of healthy and infected salmon, embedding approaches targeting a binary sentiment classification of commercial user reviews in Brazilian Portuguese, forecasting high-dimensional non-stationary time series in IoT applications, and adaptive resource allocation for communication systems. A brief description of the papers is as follows: da Silva Muniz & de Oliveira e Souza propose a new method to generate and select robust handcrafted features for music genre classification. A new feature set integrating low-and high-order statistics combined with derivatives is used, and an effective wrapper-based method for selecting feature subsets is developed, based on a consensual decision paradigm, and grounded by rigorous statistical testing.

Almeida Borges de Venancio et al. propose a hybrid method for fire detection based on spatial (visual) and temporal patterns. The proposed method works in two sequential stages: i) detection of probable fire events by a convolutional neural network based on visual patterns (spatial processing) and ii) analysis of the dynamics of these events over time (temporal processing). Experiments performed on a surveillance video database show that cascading these two stages can reduce the false positive rate with no significant impact on either the true positive rate or the processing time.

Mohiuddin et al. present a hierarchical feature selection strategy to detect deepfake images or videos. A combination of handcrafted and deep learning features is used. A hybrid feature selection method is employed that integrates a population-based meta-heuristic model, called Grey Wolf optimization, and a single solution-based meta-heuristic model, called the vortex search algorithm. The proposed model is evaluated on three publicly available datasets.

Gonzalez et al. propose a novel approach using a semi-supervised paradigm for classifying mass spectrometry data of healthy and infected salmon with the *Piscirickettsia salmonis* pathogen. The proposed algorithm based on self-training showed superior performance compared to traditional machine learning methods. Even considering a small percentage of labelled instances (25%), semi-supervised learning attains equilibrated performance across all

metrics. Furthermore, the feature selection allowed the identification of 15 potential biomarkers that define healthy and infected salmon profiles accurately.

Dias Souza & Oliveira e Souza provide a comprehensive experimental study of embedding approaches targeting a binary sentiment classification of commercial user reviews in Brazilian Portuguese. This study includes from classical (Bag-of-Words) to state-of-the-art (Transformer-based) NLP models. The fine-tuned transformer-based language models (TLMs) achieved the best results for all cases, followed by the feature-based TLM, LSTM, and CNN, with alternate ranks, depending on the database under analysis.

Bitencourt et al. propose a new methodology for forecasting high-dimensional non-stationary time series in IoT applications. A first-order MIMO multivariate model is proposed, which applies a combination of data embedding transformation and a non-stationary fuzzy time-series model to deal with concept drift events. The proposed methodology is tested on four real-world high-dimensional IoT time-series data sets.

Queiroz Carneiro et al. develop an adaptive resource allocation algorithm based on reinforcement learning for 5G multicarrier communication systems. A Markov model is introduced to describe the queueing behaviour of a communication system with 5G characteristics such as millimetre-wave propagation, high-frequency values, and specific frequency range. Then, a reinforcement learning algorithm solves the resource allocation for this communication system. The performance of the proposed adaptive resource allocation scheme is verified via computational simulations considering data packet arrivals from real traffic traces.

As guest editors of this Topical Collection, we would like to thank all the authors for their contributions, and the reviewers for their dedicated work to improve the quality of submissions. Special thanks to the Editor-in-Chief and editorial staff of the Neural Computing and Applications journal for their support in publishing this Topical Collection.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.