Special Issue:

Multi-temporal Geospatial Data Analytics and New Trends in Data and System Engineering

Guest Editors: Naphtali D. Rishe and Ivan Luković

Guest editor Naphtali D. Rishe: The first part of this issue, **Multi-temporal Geospatial Data Analytics**, is a compendium of five papers on the analysis, knowledge extraction, and application of geospatial data, including multi-band remote-sensed imagery captured by satellites over time and geolocated vector and point data.

Paper 1. **A Latent Variable Based Approach for Exploring Geographic Datasets**, by Liangdong **Deng**, Arpan Mahara, Malek Adjouadi, and Naphtali Rishe

Geographic datasets are usually accompanied by spatial nonstationarity – a phenomenon that the relationship between features varies across space. Naturally, nonstationarity can be interpreted as the underlying rule that determines how data are generated and change over space. Therefore, traditional machine learning algorithms are not suitable for handling non-stationary geographic datasets, as they only render a single global model. To solve this problem, researchers often adopt the multiple-local-model approach, which uses different models to account for different sub-regions of space. This approach has been proven efficient but not optimal, as it is inherently difficult to determine the size of subregions. Additionally, the fact that local models are only trained on a subset of data also limits their potential. This paper's authors have innovated an algorithm using an entirely different strategy that interprets nonstationarity as a lack of data and addresses it by introducing latent variables to the original dataset. Backpropagation is then used to find the best values for these latent variables. Experiments show that this method is at least as efficient as multiple-local-model-based approaches and has even greater potential.

Paper 2: **Detecting and Removing Clouds Affected Regions from Satellite Images using Deep Learning**, by Lawrence **Egharevba**, Sanjoy Kumar, Hadi Amini, Malek Adjouadi, and Naphtali Rishe

Deep Learning is becoming a very popular tool for generating and reconstructing images. Research has shown that deep learning algorithms can perform cutting-edge restoration tasks for various types of images. The performance of these algorithms can be achieved by training Deep Convolutional Neural Networks (DCNNs) with data from a large sample size. The processing of high-resolution satellite imagery becomes difficult when there are only a few images in a dataset. The authors of this paper have developed an algorithm based on the intrinsic properties of Deep Convolutional Neural Networks for the detection and removal of clouds from remote sensing images without any prior training. Their results have demonstrated that the algorithm they used performed well when compared to trained algorithms.

Paper 3: **Integrating Location Information as Geohash Codes in Convolutional Neural Network-Based Satellite Image Classification**, by Arpan **Mahara** and Naphtali Rishe

In the past few years, there have been many research studies conducted in the field of Satellite Image Classification. The purposes of these studies included flood identification, forest fire monitoring, greenery land identification, and land-usage identification. In this field, finding suitable data is often considered problematic, and some research has also been done to identify and extract suitable datasets for classification. Although satellite data can be challenging to deal with, Convolutional Neural Networks (CNNs), which consist of multiple interconnected neurons, have shown promising results when applied to satellite imagery data. In the work presented in this paper, first, the authors manually downloaded satellite images of four different classes in Florida locations using the TerraFly Mapping System. Then, the authors developed a CNN architecture suitable for extracting features and capable of multi-class classification in

their dataset. There are shortcomings in the classification due to the limited size of the dataset. To address this issue, the authors first employ data augmentation and then utilize transfer learning methodology for feature extraction with VGG16 and ResNet50 pre-trained models. The authors use these features to classify satellite imagery of Florida. The authors analyze the misclassification in their model, and to address this issue, they introduce a location-based CNN model. They convert coordinates to geohash codes, use these codes as an additional feature vector, and feed them into the CNN model. The authors contend that the new CNN model, combined with geohash codes as location features, provides better accuracy for their dataset.

Paper 4: **Spatiotemporal Model of Real Estate Valuation Trend**, by Naphtali **Rishe**, Dan Tamir, and Malek Adjouadi

The authors of this paper have developed a mathematical model and an algorithm objectivizing real estate prices so that prices across time could be compared to understand historical price trends and also to assist in a property evaluation or appraisal, as well as for the analysis of comparables, and urban/environmental changes. Given a timespan of interest, a locale (e.g., a particular zip code, a city, a county, a state), a category of properties of interest (e.g., houses), an objective historical trend in values can be computed by first evaluating the ratios between the transactions' realized prices and objective governmental assessment of the properties at some fixed point of time; then, for each period (a month) averaging the ratios of all transaction in that period; then, comparing said averages (or medians) between different periods.

Paper 5: **Towards Real-time House Detection in Aerial Imagery Using Faster Region-based Convolutional Neural Network**, by Khandaker Mamun **Ahmed**, Farid Ghareh Mohammadi, Manuel Matus, Farzan Shenavarmasouleh, Luiz Manella Pereira, Zisis Ioannis, M. Hadi Amini

In the past few years, automatic building detection in aerial images has become an emerging field in computer vision. Detecting the specific types of houses will provide information on urbanization, change detection, and urban monitoring, which play increasingly important roles in modern city planning and natural hazard preparedness. The authors of this paper have made it effective to detect various types of houses in aerial imagery using Faster Region-based Convolutional Neural Network (Faster-RCNN). After formulating the dataset and extracting bounding-box information, pre-trained ResNet50 is used to get the feature maps. The fully convolutional Region Proposal Network (RPN) first predicts the bounds and objectness score of objects (in this case, houses) from the feature maps. Then, the Region of Interest (RoI) pooling layer extracts relevant regions to detect objects that are present in the images. The authors contend that this is the first attempt at detecting houses using Faster R-CNN that has achieved satisfactory results. This work enables R&D not only in the civil and environmental domain but also in other applied science disciplines.

Naphtali David Rishe, PhD is the Inaugural Outstanding University Professor at Florida International University; Fellow of the National Academy of Inventors; Director of NSF Industry/University Cooperative Research Center for Advanced Knowledge Enablement at FIU, FAU, and Greenwich University; Director at FIU High-Performance Database Research Center; and the Director of the Geospatial Laboratory.

Guest Editor Ivan Luković: In the second part of this Special Issue **New Trends in Data and System Engineering,** we intended to address open questions and real potentials for various applications of modern approaches and technologies in data, software and system engineering in order to develop and implement effective software services in a support of information management in various organization systems. We intended to address the interdisciplinary character of a set of theories, methodologies, processes, architectures, and technologies in the development of effective software services. In a rigorous reviewing process, from more than 20 submissions, we selected the following five articles.

In the paper **A Decision Support System for Internal Migration Policy-Making**, by Boris **Delibašić**, Sandro Radovanović, and Svetlana Vukanović, a decision support system for internal migration policy in the Republic of Serbia has been proposed. The proposed system is based on machine learning and knowledge extraction methods to analyze data and identify key features for policy decision-making. The authors propose two machine learning models with an accuracy of 70% for predicting internal migration intensity in local self-governments (LSGs), as well as the decision-support tool that achieves an accuracy of 66%. The proposed system maintains desirable properties such as correctness, completeness, consistency, comprehensibility, and convenience, and allows the what-if analysis to evaluate appropriate policies for each LSG.

In the paper **An Organizational Perspective of Human Resource Modeling**, Dajana **Antanasijević**, Marko Vještica, Leposava Grubić-Nešić, Vladimir Dimitrieski, Milan Pisarić, and Sonja Ristić advocate that although Industry 4.0 improved the human-machine relationship in technical aspects, it failed to put human needs at the focus of the production process. Industry 5.0 is complementing Industry 4.0 focusing on the workers' skills, knowledge, and abilities to cooperate with machines and robots. Based on the previous research and a Domain-Specific Modeling Language (DSML) named Multi-Level Production Process Modeling Language (MultiProLan), the authors propose in the paper a new DSML named HResModLan aimed at human resource modeling from two different perspectives: organizational and production. In this paper, they present a part of HResModLan representing the organizational perspective. It is aimed at providing easier and more effective requiring, selection, hiring and development of employees within an organization. The authors present an analysis of the human resource domain, abstract and concrete syntaxes of the HResModLan language, and a model of a furniture factory and its employees expressed by the concepts of the HResModLan language.

The paper **A Meta-model for Key Performance Indicators in Higher Education** by Goran **Savić**, Milan Segedinac, Milan Čeliković, and Ivan Luković represents one of the most important results of the project PESHES – Development and implementation of a system for performance evaluation for Serbian HEIs and system, conducted through Erasmus+ program. The authors have proposed a software solution for representing diverse sets of key performance indicators in higher education. The solution addresses both the heterogeneity and the common structure of key performance indicators. To tackle the issue of heterogeneity, the authors propose a meta-model that is expressive and generic enough to represent any set of key performance indicators in higher education. The proposed solution has been verified by representing typical key performance indicator sets and developing a software application prototype that enables the creation, monitoring, and further development of key performance indicator sets.

The paper **An Analysis of Using Binary JSON Versus Native JSON on the Example of Oracle DBMS**, by Srđa **Bjeladinović**, Marko Asanović, Milica Škembarević, Olga Jejić, is about JSON, as a popular and proven standard for specifying self-describing text files with a flexible structure. To maintain its position in the market, Oracle introduced support for JSON data in the 12c R1 version of its DBMS. This version has introduced functions for storing and managing JSON data in native form. Each new version introduced new or updated JSON functions. The 21c can store JSON data in binary form. The authors in the paper present a comparison of the performances when the underlying storage of JSON is native or binary. A data model and seven use cases were designed to demonstrate earlier and new functionalities. The proposed analyses address syntax complexity, resource cost, and average execution time across functionalities for working with native and binary JSON.

The paper **Diameter-2-critical graphs of order at most 13 diameter-2 graphs** by Jovan **Radosavljević** analyses graphs that are involved in many open problems and which diameters increases by removing any of their edges (D2C graphs). A possible approach to tackle such problems is the analysis of low order graphs. For this purpose, a list of all D2C graphs of order up to 13 was obtained by incorporating the

diameter-2 test and criticality tested by *geng*, the program from the package *nauty* that generates the list of all non-isomorphic connected graphs. It seems that there are no other such results, except the previous work by the author, resulting in a list of D2C graphs of order up to10. As an application, the three conjectures concerning the maximum number of edges in D2C graphs were checked for graphs of order up to 13, and one counterexample was found.

Finally, let me express my great thanks to all the authors for their hard work, great enthusiasm, research efforts, and high-quality submissions. In the reviewing process, more than 60 reviewers were actively involved. I would like to express my warm thanks to all the reviewers for their great efforts and valuable comments that significantly contributed to raising the overall quality of the selected papers.

Ivan Luković, PhD, is a Full Professor in the Department of Information Systems of the Faculty of Organizational Sciences, University of Belgrade, Belgrade, Serbia; ORCID: 0000-0003-1319-488X and H-index = 18 (Google Scholar). His research interests are related to Database Systems, Information Systems, Business Intelligence Systems, Data Science, and Software Engineering. Chair of Managing Board of the Computer Science and Information Systems (ComSIS) journal. For several years, a chair of a series of MADEISD workshops at ADBIS conference, Track 5 – Software, System, and Service Engineering at FedCSIS conference, as well as IADSP and IADT invited sessions at KES conference. Chair of M.Sc. study program in Information Engineering, i.e., Data Science, at the Faculty of Organizational Sciences. From 2015 to 2021, created and chaired a new set of B.Sc. and M.Sc. study programs in Information Engineering, at the Faculty of Technical Sciences, University of Novi Sad. E-mail: ivan.lukovic@fon.bg.ac.rs.