Future Generation Computer Systems 88 (2018) 696–698

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Editorial

Recent advances in Big Data Analytics, Internet of Things and Machine Learning

Roshan Joy Martis [a], Varadraj Prabhu Gurupur [b], Hong Lin [c], Aminul Islam [d], Steven Lawrence Fernandes [e]

[a] Department of Electronics and Communication Engineering, Vivekananda College of Engineering and Technology, India
[b] Department of Health Management and Informatics, University of Central Florida, USA
[c] Department of Computer Science & Engineering Technology, University of Houston Downtown, USA
[d] School of Computing and Informatics, University of Louisiana at Lafayette, USA
[e] Department of Electronics and Communication Engineering, Sahyadri College of Engineering & Management, India

A B S T R A C T

Big data analytics, Internet of Things, and machine learning are some of the rising areas of science and technology forming the next generation of artificial intelligence-based computing systems. It is also important to note that this aforementioned emerging field is diverse and in some strange ways both transformative and transdisciplinary in nature. This transformative and transdisciplinary nature of this field enables it to grow both in terms of its theoretical foundations and applications.

In this special issue we focus on some advanced research projects in these areas that are transformative and transdisciplinary in nature. The projects and experiments discussed in this special issue constitute the advancement in synthesis of decision support systems that aid further advancement of healthcare delivery, diagnosing diseases, and analysis of behavioral science.

1. Introduction

Big Data Analytics, Internet of Things, and Machine Learning have a wide range of applications and combined with rich theoretical foundations. It is worthwhile noting that the theoretical foundations are transdisciplinary in nature while its applications are transformative. This transformative and transdisciplinary nature of this field [1–5] has made it one of the most coveted for scientists and innovators to explore. This exploration ranges from medical informatics to prediction of trends in the stock market. The purpose of this special issue is to navigate some of these explorations and present it to the readers. It is to be noted that while this special issue navigates these aforementioned topics it also points out important information on advancement of theoretical foundations in this field; thereby, reflecting the advancement of science.

This advancement of science is especially true when it comes to innovative healthcare systems [6]. These innovations have helped the scientific community reduce the cost of healthcare while also improving its quality. This transformation in healthcare delivery would not have been possible without the advances in theoretical science in the areas of machine learning, data mining, and knowledge networks. While machine learning and data mining have been around for a few decades the concept of knowledge networks is slowly but surely emerging as a technology disruptor especially because of its association with cloud computing. Knowledge networks and cloud computing have aided the development of diagnostic systems, tele-oncology, enabling artificial intelligence in Electronic Health Record Systems, and many other such technologies.

One of our motivations to come up with this special issue was to bring about innovations that can prove helpful in achieving the aforementioned transformative and transdisciplinary objective. One such example comes from an article published by Xie et al., [6]. Here the investigators have successfully applied advances in machine learning and Internet of Things to facilitate a transformation in healthcare delivery. Another example would be implementing smart urban planning using Big Data Analytics as described by Babar and Arif [7]. The investigators describing this article have used advances in data acquisition, data computation, and decision making formulations in developing their system. It is worthwhile noting that although the application domains for these two experiments as disparate it is to be noted that they are both transformative in nature.

Another important contribution worthwhile mentioning will be a contribution from Sarwar et al. [8] towards synthesizing a research environment for language and literature researchers. In this delineated research the investigators have introduced the techniques to manage a variety of data sources for the purpose of eliminating some of the issues associated with data access. Here we are covering an application domain that has not been covered in this section. In any case, the editors would like to mention the readers of this special issue that there is always room to transform

https://doi.org/10.1016/j.future.2018.07.057
0167-739X/

application domains using the technologies and case studies considered.

2. Contents of this special issue

In this special issue Bhushan et al., [9] have described the use of similarity measure for computing the closeness of two sets of text documents. The authors have proposed a compression modeling based on similarity measure between text documents. Two different sets of experiments were conducted to validate the efficacy of the proposed similarity measures. Experimental results demonstrate that the f-measure score obtained from similarity metric is better when compared to other text classification techniques.

Magnetic Resonance Angiography is used to detect brain images for the purpose of identifying the presence of vascular disorder. In the article titled [10], "An approach to examine magnetic resonance angiography based on Tsallis entropy and deformable snake model" the authors provide an improved approach to computerization of Magnetic Resonance Angiography using Chaotic Firefly algorithm and Tsallis entropy.

With the advent of "Internet of Things" many healthcare applications are being modernized to improve the accessibility to patients. The article titled, "Super-Resolution of Retinal Images using Multi-Kernel SVR for IoT Healthcare Applications" [11] focuses on the use of smartphone fundoscopy for the purpose diagnosing a patient remotely. This also, ties into the concept of telemedicine and telehealth.

The article titled, "Automated system for the detection of thoracolumbar fracture using a CNN architecture" [12] illustrate an automated computer aided diagnosis (CAD) system to accurately identify the fracture and prescribe correct treatment at the earliest. An automated thoracolumbar fracture detection technique is proposed using convolutional neural networks (CNNs) without segmenting the vertebra.

Bhateja, et al., [13] discuss an improvement in Computer Aided Diagnostic approach for evaluation of cancer by analyzing the mammographic masses. In their article titled, "Unsharp masking approaches for HVS based enhancement of mammographic masses: a comparative evaluation" the authors attempt to describe a method to automate the identification of the mammographic mass using image enhancements. For the purpose of experimentation, the investigators have analyzed 8 cases to validate their method. The observed qualitative measure indicates a higher performance compared to other related works.

Feature selection is an important upcoming area of research machine learning, text analysis, and image analysis. Kushwaha, and Pant [14] in their article titled, "Link-based BPSO for feature selection in big data text clustering" explain the advancement in data mining techniques using their approach. The core objective investigated is to find a new neighbor selection strategy for feature extraction. This article uses vector analysis for meeting the research objectives and provides explanation using well written algorithms.

Dogra et al., [15] have designed an image fusion technique to mitigate the artifact issues in bone and vessel. The source images are transformed using KL transformations and Ripplet transform. The artifacts are controlled via anisotropic diffusion filtering. Laplacian pyramidal based fusion is employed to fuse the mask and digital subtraction angiography images.

Tan et al., [16] describe the use of convolutional neural network in their article titled, "Age-related macular degeneration detection using deep convolutional neural network." This is another article that illustrates the use of convolutional neural networks; however, the case study this time is for problems related to ophthalmology. The core objective here is to achieve an early diagnosis of Age-related Macular Degeneration. The authors claim that their method has a higher level of accuracy compared to other available methods.

Goyal et al., [17] have described a denoising method from morphological filtering in non-sub sampled shearlet transform domain. The noisy components are processed by morphological circular disc operators i.e. Top Hat/ Bottom Hat filtering. The resultant image is decomposed into 8 bit planes and the denoised image is obtained by combining the bit planes processed with bitonic filter. Authors claim that this preserves the fine structure information when compared to other existing algorithms even at high noise levels.

The article titled, "Socio Evolution & Learning Optimization Algorithm: A Socio-inspired Optimization Methodology" [18] illustrate the use of socio-inspired algorithms in identifying mannerisms and behaviors of individuals. The authors claim that they have used 50 well-known test problems to test their algorithmic approach. This article is full of well described illustrations of the algorithms described using good examples involving good diagrams, scatter plots, and tables.

Pillai et al., [19] in their article titled "Local Diagonal Extrema Number Pattern: A new Feature Descriptor for Face Recognition" present an approach which partitions the face into several regions to facilitate extraction of features from each region individually. The extracted features are concatenated into a single feature vector which is used as a face descriptor. Only diagonal neighbours are considered, hence, the dimension of the feature, and the computation time to recognize the face are reduced.

Lin [20] described the use of Autoregressive Conditional Heteroskedastic (ARCH) and Generalized Autoregressive Conditional Heteroskedastic (GARCH) in predicting the volatility of the stock markets in China. For this the author used data pertaining to time series with associated mathematical formulation to illustrate its effective use in implementing time-varying clustering. Additionally, the author also presents variations of these models and its usage with the purpose of advancing the computational techniques involved.

Khanduzia et al., [21] in their article titled "Data envelopment analysis and interdiction median problem with fortification for enabling IoT technologies to relieve potential attacks" present a bi-objective optimization model that optimally serves customers' demands in a remarkably efficient manner and allocates fortification resources to relieve potential disaster, including interdiction operations. Authors have proposed parking management system based on IoT to assist parking drivers in managing and handling parking in the presence of protection and interdiction operations.

Acharya et al., [22] in their article titled "Automated identification of shockable and non-shockable life-threatening ventricular arrhythmias using convolutional neural network" present a tool that is used to identify shockable and non-shockable ventricular arrhythmias using neural networks. The tool described in the article uses an eleven layer convolutional neural network model. Here we perceive improvement in computational techniques to improve healthcare delivery. The editors believe that such contribution in alleviating the present techniques in computational science is much required.

The article titled, "Dynamic Metric Embedding Model for Point-of-Interest Prediction" [23] illustrate a unified flexible model which can map temporal (T) and spatial (S) factors into a unified Euclidean space and learn them collectively to resolve data sparsity. Long Short Term Memory is brought into metric embedding to depict user's dynamic interest. Furthermore, this is applied to character temporal non-uniformness. Additionally a self-attention method is developed with the ability to estimate the temporal relationship between user's behaviors.

Overall we believe that this special issue will help many future projects in the aforementioned areas of research, bolster innovative next generation computing systems, and aid creativity in the areas of big data analytics, Internet of Things, and machine