Guest Editorial Preface

Special Issue on Enabling Technologies for Smart City: IoT, Big Data, and Machine Learning

Srikanta Patnaik, SOA University, India & Interscience Institute of Management and Technology, India
Xilong Qu, School of Information Technology and Management, Hunan University of Finance and Economics, China
Tao Shen, School of Information Engineering and Automation, China
Bin Hu, Changsha Normal University, China

The concept of smart cities can only be brought to life by forming an integrated, intelligent, and interdependent network of a few core smart city technologies. When people think of smart city technologies, the words they're most likely to think of are 'IoT' and 'sensors'. While advanced sensor technology will play a critical role in empowering smart cities, a smart city is reliant on many more technologies other than sensors. These data-driven technologies function together in order to optimize resource utilization, enable convenience for citizens, and provide greater visibility and control to governments. The advances in Internet of Things and Big Data technologies as well as the availability of a wide range of machine learning algorithms offer new potentials to deliver analytical services to citizens and urban decision makers.

Smart cities, and also smart governance for that matter, rely on the collection, analysis, and processing of large volumes of granular, real-time data, which has only become possible with the help of IoT sensors. IoT sensors and cameras can constantly collect detailed information in various forms and in real-time. Data like the footfall at a railway station, traffic on a road, contamination levels in a water source, and the energy consumption in a residential block can be collected in real-time using different types of IoT sensors. By using this data, government agencies can make quick decisions pertaining to the allocation of different resources and assets. For instance, based on the footfall and ticket sales information from railway stations, transport agencies can reroute their trains to cater to the varying demand. Similarly, health, safety, and environmental agencies can monitor the contamination levels of water bodies and notify the people responsible to take remedial action. In certain cases, IoT actuators can automatically initiate responsive actions in exigent situations, such as stopping the supply of contaminated water to domestic households.

All aspects of smart cities will be primarily driven by data. All the decisions—from long-term strategic ones like public policy to short-term decisions like evaluating the welfare worthiness of individual citizens—will be taken through the analysis of pertinent data. With the increased volume, velocity, and variety of data generated thanks to IoT sensors and other advanced data gathering methods, the need for high-capacity analytics tools will be greater than ever before. Big data analytics tools are already being used by governments for a wide range of applications, from predicting the likelihood of crime in specific areas of the city to preventing crimes like child trafficking and abuse.

As IoT enables the collection of data from a host of new sources, big data analytics will be used in all areas including critical ones like education, healthcare, and transportation. Data mining is the practice of automatically searching large stores of data to discover patterns and trends that go beyond simple analysis. Data mining uses sophisticated mathematical algorithms to segment the data and evaluate the probability of future events. The process of digging through data to discover hidden connections and predict future trends has a long history. Sometimes referred to as "knowledge discovery in databases," the term "data mining" wasn't coined until the 1990s. But its foundation comprises three intertwined scientific disciplines: statistics (the numeric study of data relationships), artificial intelligence (human-like intelligence displayed by software and/or machines) and machine learning (algorithms that can learn from data to make predictions). What was old is new again, as data mining technology keeps evolving to keep pace with the limitless potential of big data and affordable computing power.

Over the last decade, advances in processing power and speed have enabled us to move beyond manual, tedious and time-consuming practices to quick, easy, and automated data analysis. The more complex the data sets collected, the more potential there is to uncover relevant insights. Retailers, banks, manufacturers, telecommunications providers, and insurers, among others, are using data mining to serve for their companies.

Machine learning is computer programs (algorithms) that can learn to do a task by doing it or by learning it from some data. For example, a computer program can be designed to learn by trying various moves and strategies against human players to select the most effective one (learning by experience). There exist many machine learning techniques for making computer programs that can learn such as artificial neural networks, support vector machines and clustering. These techniques are popular because they can be used to create programs that learn to do some complicated tasks that would be otherwise very hard to build as a computer program by a human. For example, while it is extremely hard to program a computer to recognize objects in a video by hand or to translate a text accurately, these tasks can now be learned using machine learning techniques.

The purpose of this special issue is to provide a timely opportunity for researchers, developers, engineers, students, and practitioners to share their latest discoveries in the area of data mining and machine learning and to provide a forum for them to discuss and exchange their research results, innovative ideas, and experiences in advances in data mining and machine learning, as well as to identify emerging research topics and define the future directions to achieve sustainable development. This special issue covers five (5) articles encompassing a wide range of works done in different domain pertaining to the applications of data mining and machine learning. The topics cover the data standardization of matching probability and measurement error, setting the financial development as the threshold variable and apply the dynamic panel threshold model, feature information recognition of waste recycling resource set based on Data Mining, the study on bearing capacity of autoclaved aerated concrete partition board based on deep learning, the development analysis of cross-border e-Commerce logistics based on big data technology under safety law protection.

The first article, "Smart City E-Governance Through Intelligent ICT Framework," by Dhananjaya Sarangi et al., studies how to use smart city e-governance to provide city centric services. high quality of life and the proper utilization of available resources. This article aims at explaining the Smart City project design architecture, describing how city data is collected, transmitted, stored and processed using smart IoT devices, Machine Learning (ML) and Deep Learning (DL) techniques. Through the implementation of Building Information Modelling (BIM) integrated with Six Sigma (SS) techniques, error free solutions could be provided in terms of smart applications to both citizens as well as the environment.

The second article in this issue, "Dynamic Panels with Threshold Effect of China's OFDI on Host Countries' Technological Progress: An Empirical Analysis Based on Different Types of Countries Along 'the Belt and Road'," by Sai Tan, studies and analyzes the financial development as the threshold variable and apply the dynamic panel threshold model. This paper investigates the nonlinear relationship between China's OFDI and technological progress of these countries. The results show that OFDI has different effects on technological progress of countries along the route in different thresholds ranges. Different types of countries along the route are divided into developed countries and transitional and developing countries, and the impact of transitional and developing countries is significantly greater than that of developed countries. This paper has brought important contributions: through the extended analysis based on different time periods, the author found that after China proposed the Belt and Road initiative, OFDI plays a significant role in promoting the technological progress of countries along the route. This study shows that increasing the OFDI intensity in the short term will not harm or weaken its positive technology spillover, so China should strengthen cooperation with the BRI countries to build a community of a shared future. Investment exchanges should be strengthened for the BRI countries to improve their technology-learning capacity.

Then the next article is “Feature Information Recognition of Waste Recycling Resource Set Based on Data Mining” by Qiao Jiao. The actual applicability of the existing waste recycling information identification methods is not adequate, which does not involve the basis of the algorithm, and cannot provide a reference for the sustainable development of renewable resources. For this issue, in order to improve recycling utilization and reduce waste pollution, this paper analyzes the development characteristics of waste recycling, classifies the characteristics of waste recycling, constructs the evaluation level of waste recycling feature information set, and uses data mining method to quantitatively analyze the factors affecting the development of waste recycling. This paper obtains the main factors affecting the development of renewable resources recycling and puts forward measures for developing the recycling of renewable resources and improving the overall operation supervision system of the waste chain. Experimental results show that the recovery rate of this method reaches more than 90%. After the treatment of liquid pollution, solid waste and other waste through this method, the pollution degree decreased by at least 26.3%.

The fourth article in this issue, “Study on Bearing Capacity of Autoclaved Aerated Concrete Partition Board Based on Deep Learning,” by Chunhong Zhang et al., studies the steam pressure aerated concrete board carrying capacity, puts forward construction on the basis of deep learning autoclaved aerated concrete board pressure performance research methods. Through the autoclaved aerated concrete the bearing capacity of single correlation coefficient, the relationship between the nodal force and node displacement, the calculation of the autoclaved aerated concrete stiffness, the paper obtains the autoclaved aerated concrete board yield condition. The linear buckling and the nonlinear buckling of the AUTOclaved aerated concrete sandwich panel are analyzed, and the bearing capacity of the autoclaved aerated concrete sandwich panel is calculated to realize the bearing capacity analysis. The test results show that this method can effectively improve the bearing stability of autoclaved aerated concrete sandwich.

The fifth article, “Development Analysis of Cross-Border E-Commerce Logistics Based on Big Data Technology Under Safety Law Protection,” by Wen Long, studies how to improve logistics level of cross-border e-commerce through big data technology. Firstly, difficulties existing in international logistics management and the necessity of international logistics supply chain management are discussed. Secondly, challenges and opportunities of big data technology to cross-border E-commerce are analyzed respectively. Big data logistics and its characteristics are studied in depth. The cross-border e-commerce sales forecast model is established based on GM(1,1), and the proposed model is verified based on case study, simulation results show that the prediction model can improve the prediction precision of sales amount. Finally, integration measurements of big data and cross-border E-commerce are put forward under a suitable local law protection.

The sixth article in this issue, “Research on Statistical Characteristics Modeling of Matching Probability and Measurement Error Based on Machine Learning,” by Shuan-Zhu Li et al., studies how to solve the problems that the conversion process of the current measurement error statistical characteristic model is complex, and the short-term modeling data cannot be obtained, which leads to low operation accuracy and poor processing effect, and a matching probability and measurement error statistical characteristic modeling method based on machine learning is proposed. According to the requirements of total sequence matching probability and system matching time, the sequence matching probability is calculated by machine learning, and the matching probability algorithm is optimized. Through the central limit theorem of probability theory, the sum of independent random variables is close to Gaussian distribution, the measurement error in the process of acquisition and matching is analyzed, and the measurement error is calculated for correction. Using the limited transition data, a new rough transition model process monitoring model is established to obtain the normal conversion data, online adjust and update the original model, optimize the error statistical characteristic algorithm, standardize the average value of matching measurement error, and establish the statistical model of matching probability and measurement error. The experimental results show that the average accuracy of the statistical model of matching probability and measurement error of the proposed method is 92%, which has good application effect in practical application.

I am sure the reader shall gain immense knowledge from these papers.

Srikanta Patnaik
Xilong Qu
Tao Shen
Bin Hu
Guest Editors
IJISSS