We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Chapter

Literature Review on Big Data Analytics Methods

Iman Raeesi Vanani and Setareh Majidian

Abstract

Companies and industries are faced with a huge amount of raw data, which have information and knowledge in their hidden layer. Also, the format, size, variety, and velocity of generated data bring complexity for industries to apply them in an efficient and effective way. So, complexity in data analysis and interpretation incline organizations to deploy advanced tools and techniques to overcome the difficulties of managing raw data. Big data analytics is the advanced method that has the capability for managing data. It deploys machine learning techniques and deep learning methods to benefit from gathered data. In this research, the methods of both ML and DL have been discussed, and an ML/DL deployment model for IOT data has been proposed.

Keywords: big data analytics, machine learning, deep learning, big data

1. Introduction

Digital era with its opportunity and complexity overwhelms industries and markets that are faced with a huge amount of potential information in each transaction. Being aware of the value of gathered data and benefitting from hidden knowledge create a new paradigm in this era, which redefines the meaning of power for corporation. The power of information leads organizations toward being agile and to hit the goals. Big data analytics (BDA) enforces industries to describe, diagnose, predict, prescribe, and cognate the hidden growth opportunities and leads them toward gaining business value [68]. BDA deploys advanced analytical techniques to create knowledge from exponentially increasing amount of data, which will affect the decision-making process in decreasing complexity of the process [43]. BDA needs novel and sophisticated algorithms that process and analyze real-time data and result in high-accuracy analytics. Machine and deep learning allocate their complex algorithms in this process considering the problem approach [28].

In this research, a literature review on big data analytics, deep learning and its algorithms, and machine learning and related methods has been considered. As a result, a conceptual model is provided to show the relation of the algorithms that helps researchers and practitioners in deploying BDA on IOT data.

The process of discussing over DL and ML methods has been shown in **Figure 1**.

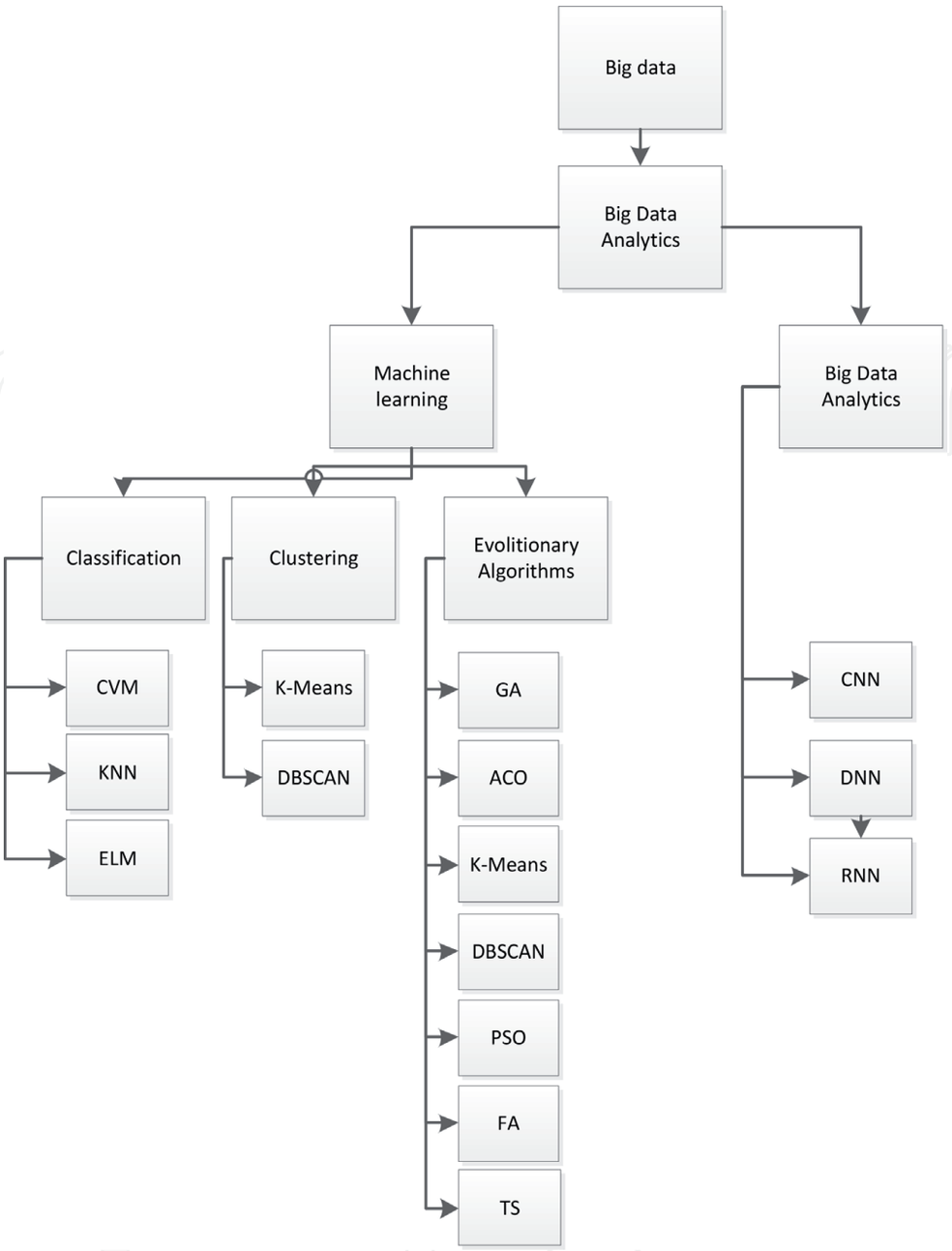


Figure 1.
The big data analytics methods in this research.

2. Big data and big data analytics

One of the vital consequences of the digital world is creating a collection of bulk of raw data. Managing such valuable capital with different shape and size on the basis of organizations' needs the manager's attention. Big data has the power to affect all parts of society from social aspect to education and all in between. As the amount of data increases especially in technology-based companies, the matter of managing raw data becomes much more important. Facing with features of raw data like variety, velocity, and volume of big data entitles advanced tools to overcome the complexity and hidden body of them. So, big data analytics has been proposed for "experimentation," "simulations," "data analysis," and "monitoring." Machine learning as one of the BDA tools creates a ground to have predictive analysis on the basis of supervised and unsupervised data input. In fact, a reciprocal relation has existed between the power of machine learning analytics and data

input; the more exact and accurate data input, the more effective the analytical performance. Also, deep learning as a subfield of machine learning is deployed to extract knowledge from hidden trends of data [28].

3. Big data analytics

In digital era with growing rate of data production, big data has been introduced, which is known by big volume, variety, veracity, velocity, and high value. It brings hardness in analyzing with itself which entitled organization to deploy a new approach and tools in analytical aspects to overcome the complexity and massiveness of different types of data (structured, semistructured, and unstructured). So, a sophisticated technique that aims to cope with complexity of big data by analyzing a huge volume of data is known as big data analytics [50]. Big data analytics for the first time was coined by Chen Chiang (2012) who pointed out the relation between business intelligence and analytics that has strong ties with data mining and statistical analysis [11].

Big data analytics supports organizations in innovation, productivity, and competition [16]. Big data analytics has been defined as techniques that are deployed to uncover hidden patterns and bring insight into interesting relations in understanding contexts by examining, processing, discovering, and exhibiting the result [69]. Complexity reduction and handling cognitive burden in knowledge-based society create a path toward gaining advantages of big data analytics. Also, the most vital feature that led big data analytics toward success is feature identification. This means that the crucial features that have important affection on results should be defined. It is followed by identifying of corelations between input and a dynamic given point, which may change during times [69].

As a result of fast evolution of big data analytics, e-business and dense connectivity globally have flourished. Governments, also, take advantages of big data analytics to serve better services to their citizens [69].

Big data in business context can be managed and analyzed through big data analytics, which is known as a specific application of this field. Also, big data gained from social media can be managed efficiently through big data analytics process. In this way, customer behavior can be understood and five features of big data, which are enumerated as volume, velocity, value, variety, and veracity, can be handled. Big data analytics not only helps business to create a comprehensive view toward consumer behavior but also helps organizations to be more innovative and effective in deploying strategies [14]. Small and medium size company use big data analytics to mine their semistructured big data, which results in better quality of product recommendation systems and improved website design [19]. As Ref. [9] cited, big data analytics gains advantages of deploying technology and techniques on their massive data to improve a firm's performance.

According to Ref. [19], the importance of big data analytics has been laid in the fact that decision-making process is supported by insight, which is the result of processing diverse data. This will turn decision-making process into an evidence-based field. Insight extraction from big data has been divided into two main processes, namely data management and data analytics with the former referring to technology support for gathering, storing, and preparing data for analyzing purpose and the latter is about techniques deployed for data analyzing and extracting knowledge from them. Thus, big data analytics has been known as a subprocess of insight extraction. Big data analytics tools are text analytics, audio analytics, video analytics, social media analytics, and predictive analytics. It can be inferred that big data analytics is the main tool for analyzing and interpreting all kinds of digital

information [35]. And the processes involved are data storage, data management, data analyzing, and data visualization [9].

Big data analytics has the potential for creating effective and efficient value in both operational and strategic approach for organization and it plays as a game changer in augmenting productivity [20].

Industry practitioners believe that big data analytics is the next 'blue ocean' that brings opportunities for organizations [33], and it is known as "the fourth paradigm of science" [70].

Fields of machine learning (ML) and deep learning (DL) were expanded to deal with BDA. Different fields like "medicine," "Internet of Things (IOT)," and "search engines" deploy ML for exploration of predictive features of big data. In other words, it generalizes learnt patterns to predict future data. Feature construction and data representation are two main elements of ML. Also, useful data extraction from big data is the reason for deploying DL, which is a human-brain inspired technique for processing neural signals as a subfield of ML [28].

4. Big data analytics and deep learning

In 1940s, deep learning was been introduced [71], but the birth of deep learning algorithms has been determined in year 2006 when layer-wise-greedy-learning method was introduced by Hinton to overcome the deficiency of neural network (NN) method in finding optimized point by trapping in optima local point that is exacerbated when the size of training data was not enough. The underlying thought of proposed method by Hinton is to use unsupervised learning before layer-by-layer training happens [72].

Inspiring from hierarchical structure of human brain, deep learning algorithms extract complex hidden features with a high level of abstraction. When massive amounts of unstructured data represent, the layered architecture of deep learning algorithms works effectively. The goal of deep learning is to deploy multiple transformation layers where in every layer output representation is occurred [42]. Big data analytics comprises the whole learnt untapped knowledge gained from deep learning. The main feature of big data analytics, which is extracting underlying features in huge amounts of data, makes it a beneficial tool for big data analytics [42].

Deep learning as a subfield of machine learning has been introduced when some conditions like rise of chip processing, which results in creating huge amounts of data, decreasing computer hardware costs, and noteworthy development in machine learning algorithms were generated. Four categories of deep learning algorithms are as follows:

- convolutional neural networks (CNN)
- restricted Boltzmann machines
- autoencoder
- sparse coding [24]

4.1 Convolutional neural networks (CNN)

CNN inspired from neural network model as a type of deep learning algorithm has a "convolutional layer" and "subsampling layer" architecture. Multi-instance data is deployed as a bag of instances in which each data point is a set of instances [73].

CNN has been known with three features namely "local field," "subsampling," and "weight sharing" and comprised of three layers, which are input, hidden that consists of "convolutional layer" and "subsampling layer" and output layer. In hidden layer, each "convolutional layer" comes after "subsampling layer." CNN training process has been done in two phases of "feed forward" in which the result of previous level entered into next level and "back propagation" pass, which is about modification of errors and deviation through a process of spreading training errors backward and in a hierarchical process [74]. In the first layer, convolution operation is deployed that is to take various filtering phases in each instances, and then, nonlinear transformation function takes place as the result of previous phase transforming into a nonlinear space. After that, the transformed nonlinear space is considered in max-pooling layer, which represents the bag of instances. This step has been done by considering the maximum response of each instance, which was in filtering step. The representation creates a strong pie with the maximum response that can be deployed by predicting instances' status in each class. This will lead to constructing a classification model [73].

CNN is comprised of feature identifier, which is an automatic learning process from extracted features from data with two components of convolutional and pooling layers. Another element of CNN is multilayer perception, which is about taking features that were learned into classification phase [3].

4.2 Deep neural network (DNN)

A deep architecture in supervised data has been introduced with advances in computation algorithm and method, which is called deep neural network (DNN) [3]. It originates from shallow artificial neural networks (SANN) that are related to artificial intelligence (AI) [30].

As hierarchical architecture of DL can constitute nonlinear information in the set of layers, DNN deploys a layered architecture with complex function to deal with complexity and high number of layers [3].

DNN is known as one of the most prominent tools for classifying [49] because of its outstanding classification performance in complex classification matters. One of the most challenging issues in DNN is training performance of it, as in optimization problems it tries to minimize an objective function with high amount of parameters in a multidimensional searching space. So, fining and training a proper DNN optimization algorithm requires in high level of attention. DNN is constructed of structure stacked denoising auto encoder (SDAE) [75] and has a number of cascade auto encoder layers and softmax classifier. The first one deploys raw data to generate novel features, and with the help of softmax, the process of feature classification is performed in an accurate way. The cited features are complementary to each other that helps DNN do its main performance, which is classification in an effective way. Gradient descent (GD) algorithm, which is an optimization method, can be deployed in linear problems with no complex objective function especially in DNN training, and the main condition of this procedure is that the amount of optimization parameter is near to optimal solution [6]. According to Ref. [30], DNN with the feature of deep architecture is deployed as a prediction model [30].

4.3 Recurrent neural network (RNN)

RNN, a network of nodes that are similar to neurons, was developed in 1980s. Each neuron-like node is interconnected with each other, and it can be divided into categories of input, hidden, and output neurons. The data will receive, transform, and generate results in this triple process. Each neuron has the feature

of time-varying real-valued activation and every synapse is real-valued weight justifiable [66]. A classifier for neural networks has outstanding performance in not only learning and approximating [105] but also in dynamic system modeling with nonlinear approach by using present data [29, 52]. RNN with the background of human brain–inspired algorithm has been derived from artificial neural network but they are slightly different from each other. Various fields of "associative memories," "image processing," "pattern recognition," "signal processing," "robotics," and "control" have been in the center of focus in research of RNN [67]. RNN with its feedback and feed forward relations can take a comprehensive view from past information and deploy it for adjusting with sudden changes. Also, RNN has the capability of using time-varying data in a recursive way, which simplified the neural network architecture. Its simplicity and dynamic features work effectively in real-time problems [40]. RNN has the ability to process temporal data in hierarchy method and take multilayer of abstract data to show dynamical features, which is another capability of RNN [18]. RNN has the potential to make connection between signals in different levels, which brings significant processing power with huge amounts of memory space [45].

5. Big data analytics and machine learning

Machine learning has been defined as predictive algorithms by data interpretation, which is followed by learning algorithm in an unstructured program. Three main categories of ML are supervised, unsupervised, and reinforcement learning [47], which is done during "data preprocessing," "learning," and "evaluation phase." Preprocessing is related to transformation of raw data into right form that can be deployed in learning phase, which comprises of some levels like cleaning the data, extracting, transforming, and combining it. In the evaluation phase, data set will be selected, and evaluation of performance, statistical tests, and estimation of errors or deviation occur. This may lead to modifying selected parameters from learning process [76]. The first one refers to analyzing features that are critical for classification through a given training data. The data deployed in training algorithm will then become trained and then it will be used in testing of unlabeled data. After interpreting unlabeled data, the output will be generated, which can be classified as discrete or regression if it is continuous. On the other hand, ML can be deployed in pattern identification without training process, which is called unsupervised ML. In this category, when pattern of characteristics are used to group the data, cluster analysis is formed, and if the hidden rules of data have been recognized, another form of ML, which is association, will be formed [77]. In the other words, the main process of unsupervised ML or clustering is to find natural grouping from those data, which is unlabeled. In this process, K cluster in a set number of data is much more similar in comparison with other clusters considering similarity measure. Three categories of unsupervised ML are "hierarchical," "partitioned," and "overlapping" techniques. "Agglomerative" and "divisive" are two kinds of hierarchical methods. The first one is referred to an element that creates a separate cluster with tendency to get involved with larger cluster; however, the second one is a comprehensive set that is going to divide into some smaller clusters. "Partitioned" methods begin with creating several disjoint clusters from data set without considering any hierarchical structure, and "overlapping" techniques are defined as methods that try to find fuzzy or deffuzy partitioning, which is done by "relaxing the mutually disjoint constraint." Among all unsupervised learning techniques, K-means grabs attention. "Simplicity" and "effectiveness" are two main characteristics of unsupervised techniques [47].

5.1 Machine learning and fuzzy logic

Fuzzy logic proposed by Lotfi Zadeh (1965) has been deployed in many fields from engineering to data analysis and all in between. Machine learning also gains advantage from fuzzy logic as fuzzy takes inductive inference. The changes happened in such grounds like "fuzzy rule induction," "fuzzy decision trees," "fuzzy nearest neighbor estimation," or "fuzzy support vector machines" [27].

5.2 Machine learning and classification methods

One of the most critical aspects of ML is classifications [23], which is the initial phase in data analytics [17]. Prior studies found new fields that can deploy this aspect like face recognition or even recognition of hand writing. According to [23], operating algorithm of classification has been divided into two categories: offline and online. In offline approach, static dataset is deployed for training. The training process will be stopped by classifiers after training process is finished and modification of data structured will not be allowed. On the other hand, online category is defined as a "one-pass" type, which is learning from new data. The prominent features of data will be stored in memory and will be kept until the processed training data is erased. Incremental and evolving processes (changing data pattern in unstable environment, which is a result of evolutionary system structure, and continuously updating meta-parameters) are two main approaches for online category [23].

Support vector machine (SVM) was proposed in 1995 by Cortes and Vapnik to solve problems related to multidimensional classification and regression issues as its outstanding learning performance [64]. In this process, SVM constructs a high-dimensional hyperplane that divides data into binary categories, and finding greatest margin in binary categories considering the hyperplane space is the main objective of this method [10]. "Statistical learning theory," "Vapnik-Chervonenkis (VC) dimension," and the "kernel method" are underlying factors of development of SVM [78], which deploys limited number of learning patterns to desirable generalization considering a risk minimization structure [22].

K-nearest neighbor deploys to classify objects in the nearest training class of features [79], and it is known as one of the most widely used algorithms in classification problems in data mining and knowledge extraction. In this method, an object is assigned to its k-nearest neighbors. The efficiency of this method is on the basis of the level of features' weighted qualifications. Some drawbacks of this method are as follows:

- It is highly dependent on the value of K parameter, which is a gauge for determination of neighborhood space.

- The method lacks discrimination ability to differentiate between far and close neighbors.

- Overlapping or noise may happen when neighbor are close [80].

KNN as one of the most important data mining algorithms was first introduced for classification problems, which are expanded to pattern recognition and machine learning research. Expert systems take advantage of KNN classification problems. Three main KNN classifiers that put focus on k-nearest vector neighbor in every class of test sample are as follows:

"Local mean-based k-nearest neighbor classifier (LMKNN)": despite the fact that existing outlier negative influence can be solved by this method, LMKNN is

prone to misclassification because of taking single value of k considering neighborhood size per class and applying it in all classes.

"Local mean-based pseudo nearest neighbor classifier (LMPNN)": LMKNN and PNN methods create LMPNN, which is known as a good classifier in "multi-local mean vectors of k-nearest neighbors and pseudo nearest neighbor based on the multi-local mean vectors for each class." Outlier points in addition to k sensitivity have been more considered in this technique. However, differentiation of information in nearest sample of classification cannot recognize widely as weight of all classes are the same [81].

"Multi-local means-based k-harmonic nearest neighbor classifier (MLMKHNN)": MLMKHNN as an extension to KNN takes harmonic mean distance for classification of decision rule. It deploys multi-local mean vectors of k-nearest neighbors per class of every query sample and harmonic mean distance will be deployed as the result of this phase [82]. These methods are designed in order to find different classification decisions [81].

In 2006, Huang et al. proposed extreme learning machine (ELM) as a classification method that works by a hidden single layer feedback in neural network [92]. In this layer, the input weight and deviation will be randomly generated and least square method will be deployed to determine output weight analytically [17], which differentiates this method from traditional methods. In this phase, learning happens followed by finding transformation matrix [93–103]. It is deployed to minimize the sum-of-squares error function. The result of minimizing function will then be used in classification or reduction of dimension [48]. Neural networks are divided into two categories of feed forward neural network and feedback neural networks and ELM is on the first category, which has a strong learning ability specially in solving nonlinear functions with high complexity. ELM uses this feature in addition to fast learning methods to solve traditional feed forward neural network problems in a mathematical change without iteration with higher speed in comparison with traditional neural network [13].

Despite the efficiency of ELM in classification problems, binary classification problems emerge as the deficiency of ELM; as in these problems, a parallel training phase on ELM is needed. In twin extreme learning machine (TELM), the problems will be solved by a simultaneous train and two nonparallel classification hyperplanes, which are deployed for classification. Every hyperplane enters into a minimization function to minimize the distance of it with one class, which is located far away from other classes [60]. ELM is at the center of attention in data stream classification research [83].

5.3 Machine learning and clustering

Clustering as a supervised learning method aims to create groups of clusters, which members of it are in common with each other in characteristics and dissimilar with other cluster members [84]. The calculated interpoint distance of every observation in a cluster is small in comparison with its distance to a point in other clusters [36]. "Exploratory pattern-analysis," "grouping," "decision-making," and "machine-learning situations" are some main applications of clustering technique. Five groups of clustering are "hierarchical clustering," "partitioning clustering," "density-based clustering," "grid-based clustering," and "model-based clustering" [84]. Clustering problems are divided into two categories: generative and discriminative approaches. The first one refers to maximizing the probability of sample generation, which is used in learning from generated models, and the other is related to deploying pairwise similarities, which maximize intercluster similarities and minimize similarities of clusters in between [63].

There are important clustering methods like K-means clustering, kernel K means, spectral clustering, and density-based clustering algorithms that are at the center of research topics for several decades. In K-means clustering, data is assigned to the nearest center, which results from being unable to detect nonspherical clusters. Kernel k-means and spectral clustering create a link between the data and feature space and after that k-means clustering is deployed. Obtaining feature space is done by using kernel function and graph model by kernel k-means and spectral clustering, respectively. Also spectral clustering deploys Eigen-decomposition techniques additionally [26]. K-means clustering works effectively in clustering of numerical data, which is multidimensional [85].

Density-based clustering is represented by DBSCAN, and clusters tend to be separate from data set and be as higher density area. This method does not deploy one cluster for clusters recognition in the data a priori. It considers user-defined parameter to create clusters, which has a bit deviation from cited parameter in clustering process [84].

5.4 Machine learning and evolutionary methods

The main goal of optimization problems is to find an optimal solution among a set of alternatives. Providing the best solution has become difficult if the searching area is large. Heuristic algorithm proposed different techniques to find the optimal solution, but they lack finding the best solution. However, population-based algorithm was generated to overcome the cited deficiency, which is considered to find the best alternative [7].

5.5 Genetic algorithms (GA)

GA is defined as a randomized search, which tries to find near-optimal solution in complex and high-dimensional environment. In GA, a bunch of genes that are called chromosomes are the main parameters in the technique. These chromosomes are deployed as a search space. A number of chromosomes that seem as a collection are called population. The creation of a random population will be followed by representing the goodness degree of objective and fitness function related to each string. The result of this step that will be a few of selected string with a number of copies will be entered into the mating pool. By deploying cross-over and mutation process, a new generation of string will be created from the string. This process will be continued until a termination condition is found. "Image processing," "neural network," and "machine learning" are some examples of application fields for genetic algorithms [38]. GA as nature-inspired algorithm is based on genetic and natural selection algorithms [31].

GA tries to find optimal solution without considering the starting point [104]; also, GA has the potential to find optimal clustering considering clustering metrics [38]. Filter and wrapper search are two main approaches of GA in the field of feature selection. The first one aims to investigate the value of features by deploying heuristic-based data characteristics like correlation, and the second one assesses the goodness of GA solution by using machine learning algorithm [53]. In K-means algorithm, optimized local point is found on the basis of initializing seed values and the generated cluster is on the basis of initial seed values. GA by the aim of finding near-optimal or optimal clustering searches for initial seed values, outperforms K-mean algorithm, and covers the lack of K-mean algorithm [4]. Gaining knowledge from data base is another ground for GA, which plays the role of building "classifier system" and "mining association rules" [58].

Feature selection is a vital problem in big data as it usually contains many features that describe target concepts and chooses proper amount of feature for

pre-processing traditionally as a main matter was done by data mining. Feature selection is divided into two groups: independent of learning algorithm, which deploys filter approach, and dependent on learning algorithm, which uses a wrapper approach. However, filter approach is independent of learning algorithm, and the optimal set of feature may be dependent on learning algorithm, which is one of the main drawbacks of filter selection. In contrast, wrapper approach by deploying learning algorithm in evaluation of every feature set works better. A main problem of this approach is complexity in computation field, which is overcome by using GA in feature selection as learning algorithm [56].

5.6 Ant colony optimization (ACO)

Ant colony optimization method was proposed by Dorigo [17] as a population-based stochastic method [15]. The method has been created biologically from real ant behavior in food-seeking pattern. In other words, this bionic algorithm has been deployed for finding the optimal path [44]. The process is that when ants start to seek food they deposit a chemical material on the ground, which is known as pheromone while they are moving toward food source. As the path between the food source and nest become shorter, the amount of pheromone will become larger. New ants in this system tend to choose the path with greater amount of pheromone. By passing time, all ants follow the positive feedback and choose the shortest path, which is signed by greatest amount of pheromone [86]. The applications of ant colony optimization in recent research have been declared as traveling salesman problem, scheduling, structural and concrete engineering, digital image processing, electrical engineering, clustering, routing optimization algorithm [41], data mining [32], robot path planning [87], and deep learning [39].

Some advantages of ant colony optimization method are as follows:

- Less complexity in integration of this method with other algorithms
- Gain advantage of distributed parallel computing (e.g., intelligent search)
- Work better in optimization in comparison with swarm intelligence
- High speed and high accuracy
- Robustness in finding a quasi-optimal solution [41]

As it is stated, the emitted material called pheromone causes clustering between species around optimal position. In big data analytics, ant colony clustering is deployed on the grid board to cluster the data objects [21].

All ant solution constructions, improvement of the movement by local search, and update of the emitted material are involved in a single iteration [23]. So, the main steps of ant colony optimization are as follows:

1. Initializing pheromone trail
2. Deploying pheromone trail to construct solution
3. Updating trail pheromone

On the basis of probabilistic state transition rule, which depends on the state of the pheromone, a complete solution is made by each ant. Two steps of evaporation

and reinforcement phase are passed in pheromone updating procedure, where evaporation of pheromone fraction happens and emitting of pheromone that shows the level of solution fitness is determined, respectively, which is followed by finalizing condition [46].

Ant colony decision tree (ACDT) is a branch of ant colony decision that aims to develop decision tress that are created in running algorithm, but as a nondeterministic algorithm in every execution, different decision tree is created. A pheromone trail on the edge and heuristics used in classical algorithm is the principle of ACDT algorithm.

The multilayered ant colony algorithm has been proposed after the disability of one layer ant colony optimization has been declared in finding optimal solution. As an item, value with massive amount of quantity takes too long to grow. In this way, through transactions, maximum quantities of an item is determined and a rough set of membership function will be set, which will be improved by refining process at subsequent levels by reduction in search space. As a result, search ranges will be differing considering the levels. Solution derived from every level is an input for next level, which is considered in the cited approach but with a smaller search space that is necessary for modifying membership functions [88]. Tsang and Kwong proposed ant colony clustering in anomaly detection [65].

5.7 Bee colony optimization (BCO)

BCO algorithm works on inspiration from honey bee's behavior, which is widely used in optimization problems like "traveling salesman problem," "internet hosting center," vehicle routing, and the list goes on. Karaboga in 2005 proposed artificial bee colony (ABC) algorithm. The main features of artificial bee colony (ABC) algorithm are simplicity, easy used and has few elements which need to be controlled in optimization problems. "Face recognition," "high-dimensional gene expression," and "speech segment classification" are some examples that ABC and ACO use to select features and optimize them by having a big search space. In ABC algorithms, three types of bees called "employed bees (EBees)," "onlooker bees (OBees)," and "scout bees deployed" are deployed. In this process, food sources are positioned and then EBees, where their numbers are equal to number of food source, pass the nectar information to OBees. They are equal to the number of EBees. The information is taken to exploit the food source till the finishing amount. Scouts in exhausted food source are employed to search for new food source. The nectar amount is a factor that shows solution quality [25, 55].

This method is comprised of two steps: step forward, which is exploring new information by bees, and step back, which is related to sharing information considering new alternative by bee of hives.

In this method, exploration is started by a bee that tries to discover a full path for its travel. When it leaves the hive, it comes across with random dances of other bees, which are equipped with movement array of other bees that is known as "preferred path." This will lead in foraging process and it comprises of a full path, which was previously discovered by its partner who guides the bee to the final destination. The process of moving from one node to another will be continued till the final destination is reached. For choosing the node by bees, a heuristic algorithm is used, which involves two factors of arc fitness and the distance heuristic. The shortest distance has the possibility to be selected by bees [7]. In BCO algorithm, two values of alpha and beta will be considered, which are exploitation and exploration processes, respectively [8].

5.8 Particle swarm optimization (PSO)

PSO was generated from inspiration from biological organisms, particularly the ability of a grouped animal to work together in order to find the desired location in

particular area. The method was introduced by Kennedy and Eberhart in 1995 as a stochastic population-based algorithm, which is known by features like trying to find global optimize point and easy implementation with taking a small amount of parameters in adjusting process. It takes benefit from a very productive searching algorithm, which makes it a best tool to work on different optimization research area and problems [59].

The searching process is led toward solving a nonlinear optimization problem in a real value search space. In this process, an iterative searching happens to find the destination, which is the optimal point. In other words, each particle has a multi-dimensional search with a specific space, which is updated by particle experience or the best neighbor's space and the objective function assesses the fitness value of each particle. The best solution, which is found in each iteration, will be kept in memory. If the optimal solution is found by particle, it is called local best or pbest and the optimal point among the particle neighbors is called global best or gbest [89]. In this algorithm, every potential solution is considered as a particle, which has several features like the current position and velocity. The balance between global and local search can be adjusted by adopting different inertia weight. One of critical success factors in PSO is a trade-off between global and local search in iteration [59]. Artificial neural network, pattern classification, and fuzzy control are some area for deploying PSO [5]. Social interaction and communication metaphor like "birds flock and fish schooling" developed this algorithm and it works on the basis of improving social information sharing, which is done among swarm particles [12].

5.9 Firefly algorithm (FA)

Firefly algorithm was been introduced by Yang [16]. The main idea of FA is that each firefly has been assumed as unisexual, which is attracted toward other firefly regardless of the gender. Brightness is the main attraction for firefly that stimulates the less bright to move toward brighter ones. The attractiveness and brightness are opposed to distance. The brightness of a firefly has been determined by the area of fitness function [90]. As the brightness of firefly increased, the level of goodness of solution increased. A full attraction model has been proposed that shows all fireflies will be attracted to brighter ones and similarity of all fireflies will occur if a great number of fireflies attract to a brighter one, which is measured by fitness value. So, convergence rate during the search method will occur in a slow pace.

FA has been inspired from the lightening feature of fireflies and known as swarm intelligence algorithm. FA better works in comparison with genetic algorithm (GA) and PSO in some cases. "Unit commitment," "energy conservation," and "complex networks" are some examples of working area of FA [61]. Fluctuation may occur when huge numbers of fireflies attract to light emission source and the searching process becomes time-consuming. To overcome these issues, neighborhood attraction FA (NaFA) is introduced, which shows that fireflies are just attracted to only some brighter points, which are outlined by previous neighbor [62].

5.10 Tabu search algorithm (TS)

Tabu search is a meta-heuristic, which was proposed by y Glover and Laguna (1997) on the basis of edge projection and making it better and it tries to make a progress in local search, which leads to a global optimized solution by taking possibility on consecutive algorithm iterations. Local heuristic search process is taken to find solution that can be deployed to combinatorial optimization paradigm [2]. The searching process in this methodology is flexible as it takes adaptive memory. The process is done during different iterations. In each iteration, a solution is found. The solution

has a neighbor point that can be reached via "move." In every move, a better solution is found, which can be stopped when no better answer is found [37]. In TS, the aspiration criteria are critical factors that lead the searching process by not considering forbidden solutions that are known by TS. In each solution, the constraints of the objective are met. So, the solutions are both feasible and time-consuming. TS process is continued by using a tabu list (TL), which is a short-term history. The short memory just keeps the recent movement, which is done by deleting the old movement when the memory is full to the maximum level [1].

The main idea of TS is to move toward solution space, which remains unexplored, which would be an opportunity to keep away from local solution. So, "tabu" movements that are recent movements are kept forbidden, which prevents from visiting previous solution points. This is proved that the method brings high-quality solutions in its iterations [57].

6. Big data analytics and Internet of Things (IOT)

Internet of things (IOT) put focus on creating an intelligent environment in which things socialize with each other by sensing, processing, communicating, and actuating activities. As IOT sensors gathered a huge amount of raw data, which is needed to be processed and analyzed, powerful tools will enforce the

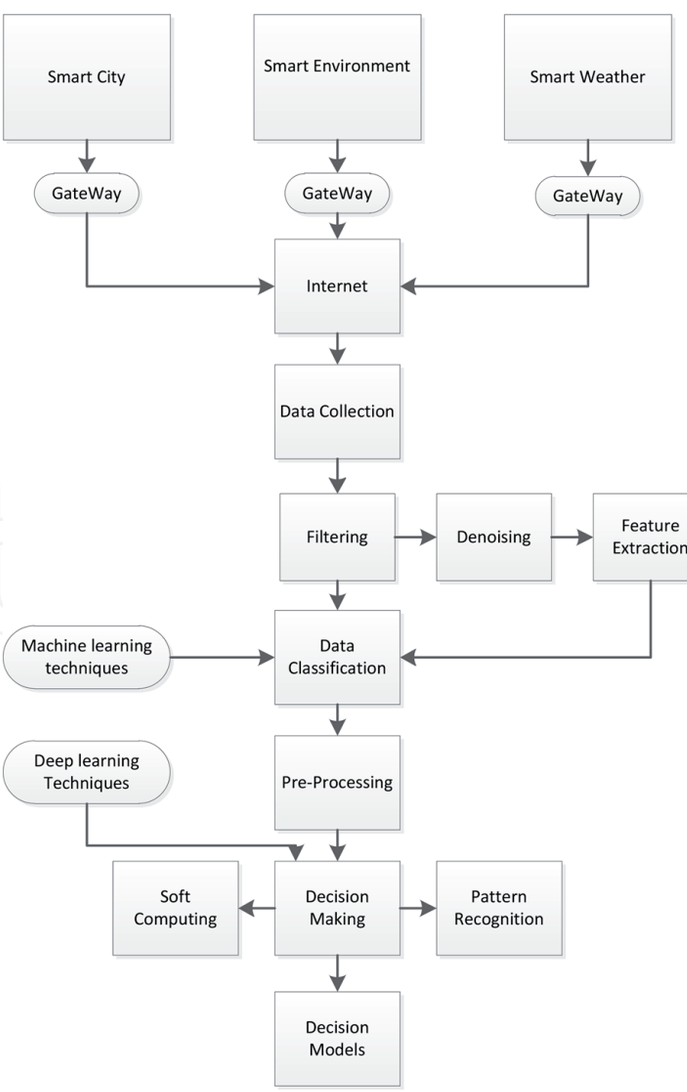


Figure 2.
IOT process.

analytics process. This will stimulate to deploy BDA and its methods on IOT-based data. Ref. [51] proposed a four-layer model to show how BDA can help IOT-based system to work better. This model comprised of data generation, sensor communication, data processing, and data interpretation [51]. It is cited that beyond 2020 cognitive processing and optimization will be considered on IOT data processing [34]. In IOT-based systems, acquired signals from sensors are gathered and deployed for processing in frame-by-frame or batch mode. Also, gathered data in IOT system will be deployed in feature extraction, which is followed by classification stage. Machine learning algorithms will be used in data classifying [54]. Machine learning classification can be deployed on three types of data, which are supervised, semisupervised, and unsupervised [54]. In decision-making level, which is comprised of pattern recognition, deep learning methods, namely, RNN, DNN, CNN, and ANN can be used for discovering knowledge. Optimization process in IOT can be used to create an optimized cluster in IOT data [91].

In **Figure 2**, the process of IOT is shown. Data is gathered from sensors. Data enters the filtering process. In this level, denoising and data cleansing happen. Also, in this level, feature extraction is considered for classification phase. After preprocessing, decision making happens on the basis of deep learning methodology (**Table 1**). Deep learning and machine learning algorithms can be used in analyzing of data generated through IOT device, especially in the classification and decision-making phase. Both supervised and unsupervised techniques can be used in classification phase considering the data type. However, both deep learning and machine learning algorithms are eligible in deploying in decision-making phase.

Phase	**Methods**		
Classification	Data type	Supervised	SVM
			Logistic regression
			Naïve Bayes
			Linear regression
			k-Nearest neighbors
		Unsupervised	Clustering
			Vector quantization
Decision-making	Deep learning methods		CNN
			RNN
			DNN
			ANN
	Machine learning optimization method		ACO
			GA
			BCO
			FFA
			PSO
			TS

Table 1.
Deep learning and machine learning techniques on IOT phases.

7. Future research directions

For feature endeavors, it is proposed to work on application of big data analytics methods on IOT fog and edge computing. It is useful to extract patterns from hidden knowledge of data gathered from sensors deploying powerful analytical tools. Fog computing is defined as a technology that is implemented in near distance to end user, which provides local processing and storage to support different devices and sensors. Health care systems gain advantage from IOT for fog computing, which supports mobility and reliability in such systems. Health care data acquisition, processing, and storage of real-time data are done in edge, cloud, and fog layer [47]. In future research, the area that machine learning algorithms can provide techniques for fog computing can be on the focus. IOT data captured from smart houses needs analytical algorithms to overcome the complexity of offline and online data gathered in processing, classification, and also next best action, or even pattern recognition [81]. Hospital information system creates "life sciences data," "clinical data," "administrative data," and "social network data." These data sources are overwhelmed with illness predictions, medical research, or even management and control of disease [39]. Big data analytics can be a future subject by helping HIS to cover data processing and disease pattern recognition.

Smart house creates ground for real-time data with high complexity, which entitles big data analytics to overcome such sophistication. Classical methods of data analyzing lost their ability in front of evolutionary methods of classification and clustering. So graphic processing unit (GPU) for machine learning and data mining purposes bring advantage for large scale dataset [7], which leads the applications into lower cost of data analytics. Another way to create future research is to work over different frameworks like Spark, which is an in-memory computation, and with the help of big data analytics, optimization problems can be solved [20].

Deployment of natural language processing (NLP) in text classification can be accompanied by different methods like CNN and RNN. These methods can gain the result with higher accuracy and lower time (Li et al., 2018).

Predictive analytics offered by big data analytics works on developing predictive models to analyze large volume data both structured and unstructured with the goal of identifying hidden patterns and relations between variables in near future [76]. Big data analytics can help cognitive computing, and behavior pattern recognition deploys deep learning technique to predict future action as it is used to predict cancer in health care system [59]. It also leads organizations to understand their problems [13].

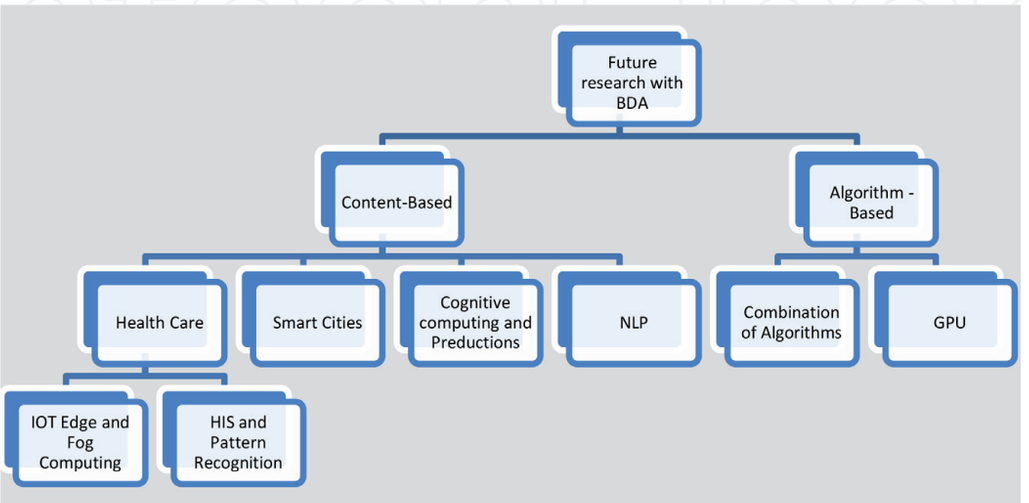


Figure 3.
Future research on big data analytics (BDA).

So, future research can be focused on both the new area for application of different machine learning or deep learning algorithm for censored data gathered and also mixture of techniques that can create globally optimal solution with higher accuracy and lower cost. Researchers can put focus on existing problems of industries through mixed application of machine learning and deep learning techniques, which may results in optimize solution with lower cost and higher speed. They also can take identified algorithms in new area of industries to solve problems, create insight, and identify hidden patterns.

In summary, future research can be done as it is shown in **Figure 3**.

8. Conclusion

This chapter has been attempted to give an overview on big data analytics and its subfields, which are machine learning and deep learning techniques. As it is cited before, big data analytics has been generated to overcome the complexity of data managing and also create and bring knowledge into organizations to empower the performances. In this chapter, DNN, RNN, and CNN have been introduced as deep learning methods, and classification, clustering, and evolutionary techniques have been overviewed. Also, a glance at some techniques of every field has been given. Also, the application of machine learning and deep learning in IOT-based data is shown in order to make IOT data analytics much more powerful in phase of classification and decision-making. It has been identified that on the basis of rapid speed of data generation through IOT sensors, big data analytics methods have been widely used for analyzing real-time data, which can solve the problem of complexity of data processing. Hospital information systems (HIS), smart cities, and smart houses take benefits of to-the-point data processing by deploying fog and cloud platforms. The methods are not only deployed to create a clear picture of clusters and classifications of data but also to create insight for future behavior by pattern recognition. A wide variety of future research has been proposed by researchers, from customer pattern recognition to predict illness like cancer and all in between are comprised in area of big data analytics algorithms.

Acknowledgements

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author details

Iman Raeesi Vanani and Setareh Majidian*
Information Technology Management, Allameh Tabataba'i University, Iran

*Address all correspondence to: setareh_majidian@atu.ac.ir

IntechOpen

© 2019 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. CC BY

References

[1] Bożejko W et al. Parallel tabu search for the cyclic job shop scheduling problem. Computers & Industrial Engineering. 2018;**113**:512-524

[2] Kiziloz H, Dokeroglu T. A robust and cooperative parallel tabu search algorithm for the maximum vertex weight clique problem. Computers & Industrial Engineering. 2018;**118**:54-66

[3] Acharya U et al. Automated detection of coronary artery disease using different durations of ECG segments with convolutional neural network. Knowledge-Based Systems. 2017;**132**:62-71

[4] Babu GP, Murty M. A near-optimal initial seed value selection in K-means algorithm using a genetic algorithm. Pattern Recognition Letters. 1993;**14**(10):763-769

[5] Bonyadi MR, Michalewicz Z. Particle swarm optimization for single objective continuous space problems: A review. Evolutionary Computation. 2017;**25**(1):1-54

[6] Caliskan A et al. Classification of high resolution hyperspectral remote sensing data using deep neural networks. Engineering Applications of Artificial Intelligence. 2018;**67**:14-23

[7] Cano A. A survey on graphic processing unit computing for large-scale data mining. WIREs Data Mining and Knowledge Discovery. 2017;**8**(1):e1232. DOI: 10.1002/widm.1232

[8] Caraveo C et al. Optimization of fuzzy controller design using a new bee colony algorithm with fuzzy dynamic parameter adaptation. Applied Soft Computing. 2016;**43**:131-142

[9] Castillo O, Amador-Angulo L. A generalized type-2 fuzzy logic approach for dynamic parameter adaptation in bee colony optimization applied to fuzzy controller design. Information Sciences. 2018;**460**-**461**:476-496

[10] Chen J et al. The synergistic effects of IT-enabled resources on organizational capabilities and firmperformance. Information and Management. 2012;**49**(34):140-152

[11] Chou J et al. Metaheuristic optimization within machine learning-based classification system for early warnings related to geotechnical problems. Automation in Construction. 2016;**68**:65-80

[12] Côrte-Real A et al. Assessing business value of Big Data Analytics in European firms. Journal of Business Research. 2017;**70**:379-390

[13] Côrte-Real N et al. Unlocking the drivers of big data analytics value in firms. Journal of Business Research. 2019;**97**:160-173

[14] Delice Y et al. A modified particle swarm optimization algorithm to mixed-model two-sided assembly line balancing. Journal of Intelligent Manufacturing. 2017;**28**(1):23-36

[15] Ding S et al. Extreme learning machine: Algorithm, theory and applications. Artificial Intelligent Review. 2015;**44**(1):103-115

[16] Dong J, Yang C. Business value of big data analytics: A systems-theoretic approach and empirical test. In: Information & Management. 2018. [In Press]

[17] Dorigo M. Ant Colony Optimization: New Optimization Techniques in Engineering. Berlin Heidelberg: Springer-Verlag; 1991. pp. 101-117

[18] Esposito C et al. A knowledge-based platform for Big Data analytics based on publish/subscribe services and stream processing. Knowledge-Based Systems. 2015;**79**:3-17

[19] Feng L et al. Rough extreme learning machine: A new classification method based on uncertainty measure. Neurocomputing. 2019;**325**:269-282

[20] Gonzalez-Lopez J et al. Distributed nearest neighbor classification for large-scale multi-label data on spark. Future Generation Computer Systems. 2018;**87**:66-82

[21] Gallicchio C et al. Deep reservoir computing: A critical experimental analysis. Neurocomputing. 2017;**268**:87-99

[22] Gandomi A, Haider M. Beyond the hype: Big data concepts, methods, and analytics. International Journal of Information Management. 2015;**35**(2):137-144

[23] German F et al. Do retailers benefit from deploying customer analytics? Journal of Retailing. 2014;**90**:587-593

[24] Ghosh A et al. Aggregation pheromone density based data clustering. Information Sciences. 2008;**178**:2816-2831

[25] Gonzalez-Abril L et al. Handling binary classification problems with a priority class by using support vector machines. Applied Soft Computing. 2017;**61**:661-669

[26] Gu X, Angelov P. Self-organizing fuzzy logic classifier. Information Sciences. 2018;**447**:36-51

[27] Guo Y et al. Deep learning for visual understanding: A review. Neurocomputing. 2016;**187**:27-48

[28] Harfouchi F et al. Modified multiple search cooperative foraging strategy for improved artificial bee colony optimization with robustness analysis. Soft Computing. 2017;**22**(19)

[29] Huang J et al. A clustering method based on extreme learning machine. Neurocomputing. 2018;**227**:108-119

[30] Hüllermeier E. Does machine learning need fuzzy logic? Fuzzy Sets and Systems. 2015;**281**:292-299

[31] Jan B et al. Deep learning in big data analytics: A comparative study. Computers and Electrical Engineering. 2017;**75**:1-13

[32] Jiang P, Chen J. Displacement prediction of landslide based on generalized regression neural networks with K-fold cross-validation. Neurocomputing. 2016;**198**:40-47

[33] Jiang S et al. Modified genetic algorithm-based feature selection combined with pre-trained deep neural network for demand forecasting in outpatient department. Expert Systems With Applications. 2017;**82**:216-230

[34] Ko Y. How to use negative class information for Naive Bayes classification. Information Processing and Management. 2017;**53**(6):1255-1268

[35] Koonce D, Tsaib S. Using data mining to find patterns in genetic algorithm solutions to a job shop schedule. Computers & Industrial Engineering. 2000;**38**(3):361-374

[36] Kozak J, Boryczka U. Collective data mining in the ant colony decision tree approach. Information Sciences. 2016;**372**:126-147

[37] Kwon O et al. Data quality management, data usage experience and acquisition intention of big data analytics. International Journal of Information Management. 2014;**34**(3):387-394

[38] Lee I, Lee K. The Internet of Things (IoT): Applications, investments, and challenges for enterprises. Business Horizons. 2015;**58**(4):1-10

[39] Li J et al. Medical big data analysis in hospital information system. In: Big Data on Real-World Applications. 2016. Chapter 4

[40] Loebbecke C, Picot A. Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. Journal of Strategic Information Systems. 2015;**24**(3):149-157

[41] Lohrmann C, Luukka P. A novel similarity classifier with multiple ideal vectors based on k-means clustering. Decision Support Systems. 2018;**111**:27-37

[42] Martí R et al. Tabu search for the dynamic bipartite drawing problem. Computers and Operations Research. 2018;**91**:1-12

[43] Maulik U et al. Genetic algorithm-based clustering technique. Pattern Recognition. 2000;**33**(9):1455-1465

[44] Mavrovounioti M, Yang S. Training neural networks with ant colony optimization algorithms for pattern classification. Journal of Soft Computing. 2015;**19**(6):1511-1522

[45] Miao Z et al. Robust tracking control of uncertain dynamic nonholonomic systems using recurrent neural networks. Neurocomputing. 2014;**142**:216-227

[46] Mohan B, Baskaran R. A survey: Ant colony optimization based recent research and implementation on several engineering domain. Expert Systems with Applications. 2012;**39**(4):4618-4627

[47] Mutlag AA et al. Enabling technologies for fog computing in health care IoT systems. Future Generation Computer Systems. 2019;**90**:62-78

[48] Najafabadi M et al. Deep learning applications and challenges in big data analytics. Journal of Big Data. 2015:1-21. DOI: 10.1186/s40537-014-0007-7

[49] Nguyen T et al. Big data analytics in supply chain management: A state-of-the-art literature review. Computers and Operations Research. 2018;**98**:254-264

[50] Ning J et al. A best-path-updating information-guided ant colony optimization algorithm. Information Sciences. 2018;**433-434**:142-162

[51] Osipov V, Osipova M. Space–time signal binding in recurrent neural networks with controlled elements. Neurocomputing. 2018;**308**:194-204

[52] Panda M, Abraham A. Hybrid evolutionary algorithms for classification data mining. In: Neural Computing & Applications. 2014;**26**(3):507-523

[53] Peng H et al. An unsupervised learning algorithm for membrane computing. Information Sciences. 2015;**304**:80-91

[54] Peng Y et al. Orthogonal extreme learning machine for image classification. Neurocomputing. 2017;**266**:458-464

[55] Qawaqneh Z et al. Age and gender classification from speech and face images by jointly fine-tuned deep neural networks. Expert Systems With Applications. 2017;**85**:78-86

[56] Ramsingh J, Bhuvaneswari V. An efficient map reduce-based hybrid NBC-TFIDF algorithm to mine the public sentiment on diabetes mellitus—A big data approach. Journal of King Saud University – Computer and Information Sciences. 2018. [In Press]

[57] Rathore M et al. Urban planning and building smart cities based on the Internet

of Things using Big Data analytics. Computer Networks. 2016;**101**:63-80

[58] Ruan X, Zhang Y. Blind sequence estimation of MPSK signals using dynamically driven recurrent neural networks. Neurocomputing. 2014;**129**:421-427

[59] Sekaran K et al. Deep learning convolutional neural network (CNN) with Gaussian mixture model for predicting pancreatic cancer. Multimedia Tools and Applications. 2019:1-15. DOI: 10.1007/s11042-019-7419-5

[60] Shah S, Kusiak A. Data mining and genetic algorithm based gene/SNP selection. Artificial Intelligence in Medicine. 2004;**31**(3):183-196

[61] Shanthamallu U et al. A brief survey of machine learning methods and their sensor and IoT applications. In: 2017 8th International Conference on Information, Intelligence, Systems & Applications (IISA). 2017. DOI: 10.1109/IISA.2017.8316459

[62] Shunmugapriya P, Kanmani S. A hybrid algorithm using ant and bee colony optimization for feature selection and classification (AC-ABC Hybrid). Swarm and Evolutionary Computation. 2017;**36**:27-36

[63] Sikora R, Piramuthu S. Framework for efficient feature selection in genetic algorithm based data mining. European Journal of Operational Research. 2007;**180**(2):723-737

[64] Silva M, Cunha C. A tabu search heuristic for the uncapacitated single allocation p-hub maximal covering problem. European Journal of Operational Research. 2017;**262**(3):954-965

[65] Srinivasa KG et al. A self-adaptive migration model genetic algorithm for data mining applications. Information Sciences. 2007;**177**(20):4295-4313

[66] Taherkhani M, Safabakhsh R. A novel stability-based adaptive inertia weight for particle swarm optimization. Applied Soft Computing. 2016;**38**:281-295

[67] Wan Y et al. Twin extreme learning machines for pattern classification. Neurocomputing. 2017;**260**:235-244

[68] Wang H et al. Firefly algorithm with neighborhood attraction. Information Sciences. 2017;**382-383**:374-387

[69] Wang H et al. Randomly attracted firefly algorithm with neighborhood search and dynamic parameter adjustment mechanism. Journal of Soft Computing. 2017;**21**(18):5325-5339

[70] Wang Q et al. Local kernel alignment based multi-view clustering using extreme learning machine. Neurocomputing. 2018;**275**:1099-1111

[71] Wu J et al. A patent quality analysis and classification system using self-organizing maps with support vector machine. Applied Soft Computing. 2016;**41**:305-316

[72] Zhang L, Zhang Q. A novel ant-based clustering algorithm using the kernel method. Information Sciences. 2011;**181**:4658-4672

[73] Zhang X et al. An overview of recent developments in Lyapunov–Krasovskii functionals and stability criteria for recurrent neural networks with time-varying delays. Neurocomputing. 2018;**313**:392-401

[74] Zhu S, Shen Y. Robustness analysis for connection weight matrix of global exponential stability recurrent neural networks. Neurocomputing. 2013;**101**:370-374

[75] Wang Y et al. Integrated big data analytics-enabled transformation model: Application to health care. Information and Management. 2018;**55**(1):64-79

[76] Wang Y, Hajli N. Exploring the path to big data analytics success in healthcare. Journal of Business Research. 2017;**70**:287-299

[77] Iqbal R et al. Big data analytics: Computational intelligence techniques and application areas. Technological Forecasting & Social Change. 2018. [In Press]

[78] Wamba S et al. Big data analytics and firm performance: Effects of dynamic capabilities. Journal of Business Research. 2017;**70**:356-365

[79] Zhang Q et al. A survey on deep learning for big data. Information Fusion. 2018;**42**:146-157

[80] Liu W et al. A survey of deep neural network architectures and their applications. Neurocomputing. 2017;**234**:11-26

[81] Yassine A et al. IoT big data analytics for smart homes with fog and cloud computing. Future Generation Computer Systems. 2019;**91**:563-573

[82] Yin Z et al. A-optimal convolutional neural network. Neural Computings & Applications. 2016;**30**(7):2295-2304

[83] Wang S et al. Convolutional neural network-based hidden Markov models for rolling element bearing fault identification. Knowledge-Based Systems. 2018;**144**:65-76

[84] Shi X et al. Tracking topology structure adaptively with deep neural networks. Neural Computing & Application. 2017;**30**(11):3317-3326

[85] Zhou L et al. Machine learning on big data: Opportunities and challenges. Neurocomputing Journal. 2017;**237**:350-361

[86] Tack C. Artificial intelligence and machine learning | applications in musculoskeletal physiotherapy. Musculoskeletal Science and Practice. 2018;**39**:164-169

[87] Tang L et al. A novel perspective on multiclass classification: Regular simplex support vector machine. Information Sciences. 2018;**480**:324-338

[88] Xia M et al. A hybrid method based on extreme learning machine and k-nearest neighbor for cloud classification of ground-based visible cloud image. Neurocomputing. 2015;**160**:238-249

[89] Onan A. A fuzzy-rough nearest neighbor classifier combined with consistency-based subset evaluation and instance selection for automated diagnosis of breast cancer. Expert Systems with Applications. 2015;**42**(20):6844-6852

[90] Gou J et al. A generalized mean distance-based k-nearest neighbor classifier. Expert Systems With Applications. 2019;**115**:356-372

[91] Pan Z et al. A new k-harmonic nearest neighbor classifier based on the multi-local means. Expert Systems With Applications. 2017;**67**:115-125

[92] Xu S, Wang J. Dynamic extreme learning machine for data stream classification. Neurocomputing. 2017;**238**:433-449

[93] Du G et al. Study on density peaks clustering based on k-nearest neighbors and principal component analysis. Knowledge-Based Systems. 2016;**99**:135-145

[94] Yu S et al. Two improved k-means algorithms. Applied Soft Computing. 2018;**68**:747-755

[95] Tabakhi S et al. Gene selection for microarray data classification using a novel ant colony optimization. Neurocomputing. 2015;**168**:1024-1036

[96] Liu H et al. A path planning approach for crowd evacuation in buildings based on improved artificial bee colony algorithm. Applied Soft Computing. 2018;**68**:360-376

[97] Hong T et al. A multi-level ant-colony mining algorithm for membership functions. Information Sciences. 2012;**182**(1):3-14

[98] Kuo RJ et al. Integration of growing self-organizing map and bee colony optimization algorithm for part clustering. Computers & Industrial Engineering. 2018;**120**:251-265

[99] Verma O et al. Opposition and dimensional based modified firefly algorithm. Expert Systems With Applications. 2016;**44**:168-176

[100] Janakiraman S. A hybrid ant colony and artificial bee colony optimization algorithm-based cluster head selection for IoT. Procedia Computer Science. 2018;**143**:360-366

[101] Tsai C et al. Metaheuristic algorithms for healthcare: Open issues and challenges. Computers and Electrical Engineering. 2016;**53**:421-434

[102] Villarrubia G et al. Artificial neural networks used in optimization problems. Neurocomputing. 2018;**272**:10-16

[103] Wari E, Zhu W. A survey on metaheuristics for optimization in food manufacturing. Applied Soft Computing. 2016;**46**:328-343

[104] Wu J et al. Evolving RBF neural networks for rainfall prediction using hybrid particle swarm optimization and genetic algorithm. Neurocomputing. 2015;**148**:136-142

[105] Yang F et al. A new approach to non-fragile state estimation for continuous neural networks with time-delays. Neurocomputing. 2016;**197**:205-211