Research Article

A New Random Forest Algorithm Based on Learning Automata

Mohammad Savargiv [ID],[1] **Behrooz Masoumi** [ID],[1] **and Mohammad Reza Keyvanpour**[2]

[1]Faculty of Computer and Information Technology Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran
[2]Department of Computer Engineering, Alzahra University, Tehran, Iran

Correspondence should be addressed to Behrooz Masoumi; masoumi@qiau.ac.ir

The goal of aggregating the base classifiers is to achieve an aggregated classifier that has a higher resolution than individual classifiers. Random forest is one of the types of ensemble learning methods that have been considered more than other ensemble learning methods due to its simple structure, ease of understanding, as well as higher efficiency than similar methods. The ability and efficiency of classical methods are always influenced by the data. The capabilities of independence from the data domain, and the ability to adapt to problem space conditions, are the most challenging issues about the different types of classifiers. In this paper, a method based on learning automata is presented, through which the adaptive capabilities of the problem space, as well as the independence of the data domain, are added to the random forest to increase its efficiency. Using the idea of reinforcement learning in the random forest has made it possible to address issues with data that have a dynamic behaviour. Dynamic behaviour refers to the variability in the behaviour of a data sample in different domains. Therefore, to evaluate the proposed method, and to create an environment with dynamic behaviour, different domains of data have been considered. In the proposed method, the idea is added to the random forest using learning automata. The reason for this choice is the simple structure of the learning automata and the compatibility of the learning automata with the problem space. The evaluation results confirm the improvement of random forest efficiency.

1. Introduction

Random forest is one of the methods of ensemble learning that comes under the homogeneous base learner category in terms of the type of constructive classifiers. As the name implies, all base learners are decision trees, and therefore they have a simpler structure than similar methods [1]. The random forest structure has two advantages. The first category is from a computational point of view, and the second category is from a statistical point of view. Advantages that can be considered from a computational point of view are: the random forest has the ability to deal with both regression and classification issues. The train and prediction processes in this classifier are performed at high speed, and therefore the random forest is known as one of the fast classic classifiers. Another advantage of the random forest is its ability to be used directly in high-dimensional issues [2]. The advantages of the second view of the random forest are its

characteristics, namely, prioritization of features, attribution of different weight coefficients to different classes, and illustration and unsupervised learning ability.

According to the literature, the random forest method is one of the most practical methods of ensemble learning. Weighting the base learners in ensemble learning is one of the main challenges in aggregating the basic classifiers in order to achieve a stronger classifier [3]. The reason for weighing base learners, or in other words, determining the impact factor for each base learner, is to increase the scalability of the data mining algorithm with the problem space. This becomes even more apparent when the environment is dynamic, and different or sometimes contradictory behaviours are observed from data in different situations. The text data environment has such an interesting behaviour that it challenges data mining algorithms. For example, placing one word on one domain may create a positive polarity, but it may also create a negative polarity on another domain. This

difference in polarity is created without any change in the form of the word and without any change in the role of the word from a grammatical point of view. The word "small" in both the electronic domain and the restaurant domain has such a behaviour. This behaviour poses a major challenge to the opinion mining algorithms [4].

The classical solution in the literature to overcome this challenge is based on the use of lexical-based approaches. This approach is based on frameworks such as unigram, n-gram, aspect-based, and similar methods, and all of them are data-dependent. In addition to the urgent need for predefined data, these methods lose their efficiency if they are met with an unspecified word or metaphor in the opinion mining field. In other words, they are not compatible with the problem space. The way random forest works is that with the sequential placement of training data and feature vectors that are injected into each of the base learners, it tries to find the best subset of features, and by increasing their impact factor in the classifier, it achieves the highest performance among all the aggregated base learners [5]. However, this method is not effective in relation to data such as text, in which a word can have different polarities in different domains because, in the classification algorithm, there is no ability to adapt to the conditions of the problem space.

In this paper, we intend to empower random forest with the idea of reinforcement learning and improve its efficiency. In the proposed method, learning automata is used to aggregate and weigh base learners. The way learning automata works is to receive feedback from the environment and perform one of the actions based on the type of feedback. In the learning automata, feedbacks are divided into two categories of reinforcement signals: reward signals and penalty signals. For each reinforcement signal received by the learning automata, it updates the probability of selecting the selected action in the previous step. This process continues until the probability of action selections converges to one of the actions; in other words, the best option for running in the current situation is found. In the proposed method, learning automata actions are appropriate when one of the base learners selected leads to the maximum reward that can be received from the environment. Since at each stage of learning automata execution, the learning algorithm tries to select the best option, achieving global optima in the problem space is guaranteed. This is proof of the adaptability of the proposed method. In the proposed method, the subprocess of replacing features in the feature vector is removed, and all the features in the feature vector are used. As a practical application in the field of opinion mining, if the Bag of Word (BoW) method is used to create the feature vector, the advantage of considering all the features of the feature vector will also cover cases that occur rarely. In other words, in the proposed method, the aspect of independence from the domain in the processes such as opinion mining is considered.

Our contribution is summarized as follows:

In this paper, a brief review of random forest in terms of application scope is given.

In this paper, a learning automata-based method is proposed to improve the random forest performance.

The proposed method operates independently of the domain, and it is adaptable to the conditions of the problem space.

The rest of the paper is organized as follows. In Section 2, related work is introduced. Section 3 presents the introduction to learning automata. The proposed method is explained in Section 4. Section 5 includes evaluation. Discussion is given in Section 6, and finally, the conclusion and future work are described in Section 7.

2. Related Work

In this section, theories and literature on the subject of random forest are examined. The purpose of this section is to review the innovations that have been introduced around random forest in recent years.

Random forest is considered as one of the methods of ensemble learning in the homogeneous ensemble learning subgroup. In the random forest, each decision tree, or in other words, each base learner, has access to a random subset of feature vectors [6]. Therefore, the feature vector is defined as follows:

$$x = (x_1, x_2, ..., x_p), \tag{1}$$

, where p is the dimension property of the available vector for the base learner. The main goal is to find the prediction function as $f(x)$ that predicts the Y parameter. The prediction function is defined as follows:

$$L(Y, f(x)), \tag{2}$$

where L is known as the loss function, and the goal is to minimize the expected value of the loss. For regression applications and classification applications, squared error loss and zero-one loss are common choices, respectively. These two functions are defined as follows in equations (3) and (4), respectively.

$$L(Y, f(x)) = (Y - f(x)^2), \tag{3}$$

$$L(Y, f(x)) = I(Y \neq f(x)) = \begin{cases} 0, \text{ if } Y = f(x), \\ 1, \text{ otherwise.} \end{cases} \tag{4}$$

To create an ensemble, a set of base learners come together. If base learners are defined as follows:

$$h_1(x), h_2(x), \ldots, h_J(x), \tag{5}$$

for regression applications, the averaging will be based on equation (6), and for classification applications, the voting will be based on equation (7).

$$f(x) = \frac{1}{J} \sum_{j=1}^{J} h_j(x), \qquad (6)$$

$$f(x) = \arg\max \sum_{j=1}^{J} I(y = h_j(x)). \qquad (7)$$

The Random Forest pseudocode for classification applications is shown in Algorithm 1.

As can be seen in Algorithm 1, in the random forest, an attempt is made to find a subset of features using the various replacements of training data and features that maximize the efficiency and accuracy of the output. This set of features is used to identify a new instance.

The following is a brief review of the random forest subject literature. It should be noted that we intend to introduce the background of the subject, and this paper is not a review paper, and the presented review is a brief review and does not mention all the previous works undoubtedly. However, the authors have tried to refer to the latest and most authoritative research work published in the recent years.

2.1. Astronomy, Bioinformatics, and Economics fields.

In the astronomy field, Markel and Bayless [7] use RF for the classification of type IA and core-collapse supernovae. Chen et al. [8] propose an approach to detect the potential signal photons by RF. In the bioinformatics, Pang et al. [9] propose a method to mitigate the computational complexity of RNA simulation software by a typical random forest. Darmawan et al. [10] propose an age estimation model in the bioinformatics field. In the economics field, Park et al. [11] propose two stages of short-term load forecasting by random forest and deep neural networks to reduce energy costs. 12 use a typical RF to solve the e-commerce product classification problem. Modeling consumer credit risk by RF is the main goal of [13]. 14 increase tree correlation by controlling the probability of placing splits along with strong predictors to deal with high-dimensional settings. Sikdar et al. [15] proposed a variable selection method based on RF to identify the key predictors of price change in amazon.

2.2. General and Global Problem fields.

In the general field, Giffon et al. [16] use the mean of orthogonal matching pursuit algorithms for calculating the weights of the linear combination for producing a linear combination of trees with minimum training error. Combining RF and generalized linear mixed models is the main idea of [17] to model clustered and longitudinal binary outcomes. Mohapatra et al. [18] optimize the random forest by use of unequal weight voting strategy. Ji et al. [19] propose a hybrid model for crowd counting by a combination of convolutional neural networks (CNN) and deep regression forest. Santra et al. [20] propose a deterministic dropout to remove unimportant connections in NN by RF. Proposing the oblique RF without explicit regularization techniques by minimizing the structural risk is the main goal of [21]. Katuwal et al. [22]

use an oblique hyperplane to split the data for increasing the accuracy of the trees and reduce the depth of RF. Probst et al. [23] tune the hyper-parameters to achieve higher performance to improve the RF. Kim et al. [24] propose a method for interpreting and simplifying a black-box model of a deep RF by quantifying the feature contributions and frequency of the fully trained deep RF. Jain et al. [25] purpose dynamic weighing scheme for RF using the correlation between decision tree and data samples. In the global problem field, Stafoggia et al. [26] estimate daily particulate matter for weather forecasting by RF. Modeling the global forest area by RF is the main target of [27]. Breidenbach and Saravi [28] present research on land-subsidence spatial modeling and its assessment. Analyzing the net ecosystem carbon exchange is the goal of [29]. Prediction about the global climate problem using the index quantization ability of random forest and the optimizing ability of PSO in the NN prediction model is the main purpose of [30]. Li et al. [31] solve the class imbalance by detecting serial case pairs.

2.3. Healthcare field.

Diagnosis detection and prediction of obesity in patients by RF are the main goals of [32, 33], respectively. El-Sappagh et al. [34] use RF in the simple form for the detection of Alzheimer's disease progression. In [35], RF is introduced as one useful machine learning tool for healthcare domain, especially for COVID-19 modeling. Khedkar et al. [36] use Patients Electronic Health Records for predicting the heart failure risks by RF. Hane et al. [37] propose a model for prediction of the dissolution behaviour of a wide variety of oxide glasses. Subudhi et al. [38] propose a method by RF to detect the ischemic stroke by a sequence of MRI images. Javadi et al. [39] propose a method to predict the immunogenic peptides of intracellular parasites. Identifying the key risk factors associated with acute rejection in organ transplantation is the main propose of [40]. In Singh et al. [41], RF has been used as one of the classifiers to classify the covid-19 spread. Na et al. [42] propose an automatic walking mode change of the above-knee prosthesis. Clustering and predicting vital signs by RF is the goal of [43]. Zhu et al. [44] optimize the parameters of the random forest by improved fish Swarm algorithm for predicting the knee contact force. A method for identifying foreign particles for quality detection of liquid pharmaceutical products is presented by [45]. Lee and Jung [46] consider the relation between teacher attachment and student growth. 47 propose a practical method for SIF downscaling. Guanter et al. [48] present a method based on RF for predicting diabetes. Subasi et al. [49] propose a decision support system for the diagnosis of migraine by RF. Classification of the driver's stress level is the main goal of [50]. Ayata et al. [51] propose an emotion recognition algorithm from multimodal physiological signals by using the random forest as one of the machine learning methods for recognition.

2.4. Industrial and Network fields.

Zeraatpisheh et al. [52] use typical RF for producing the feature map in the industrial field. Du et al. [53] propose a rapid and accurate detection technique for pesticide detection by RF to

Let $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$ denote the training data, with $x_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,p})^{\mathrm{T}}$
For $j = 1$ to J:
Take a bootstrap sample D of size N from D.
Using the bootstrap sample, D_j as the training data fit a tree.
(a)   Start with all observations in a single node.
(b)   Repeat the following steps recursively for each node until the stopping criterion is met: (i) Select m predictors at random from the p available predictors.
Find the best binary split among all binary splits in the predictors from step (i).
Split the node into two descendant nodes using the split from step (ii).
To make a prediction at a new point x.
$\widehat{f}(x) = \mathrm{argmax}_y \sum_{j=1}^{J} I(\widehat{h}_j(x))$
Where $\widehat{h}_j(x)$ is the prediction of the response variable at x using the jth tree.

ALGORITHM 1: The random forest pseudocode for classification applications [1].

construct a quantitative detection model. Improving the performance of mapping for mineral is the main goal of reference [54]. Liu et al. [55] propose an adaptive electrical period partition algorithm for open-circuit fault detection. Software fault prediction by ensemble techniques is investigated by [56]. In [57], the RF id is used to build a distributed energy system. A comprehensive image processing model is proposed by [58]. Ho et al. [59] uses RF to propose a framework that uses climate data to model hydropower generation. Zhou et al. [60] use RF for small and unbalanced datasets to create a risk prediction model for decision-making tool. Deng et al. [61] propose an authentication method for protecting high-value food products by RF. The forecast for agricultural products by RF is proposed by [62]. Jeong and Kim [63] use weighted random forest for the link prediction model. Khorshidpour et al. [64] present an approach to model an attack against classifiers with non-differentiable decision boundary. Fusing multi-domain entropy and RF is the main goal of [65] for proposing a fault diagnosis method of the inter-shaft bearing. Analyzing the wine quality is presented by [66]. In the network field, Madhumathi and Suresh [67] develop a model to predict the future location of a dynamic sensor node in wireless communications. Fang et al. [68] propose an encrypted malicious traffic identification method. Detecting the intrusion in the network by typical RF is proposed by [69], and intrusion detection in the network security by tuning the RF parameter of the Moth-Flame optimization algorithm is presented by [70].

2.5. Physics, Text Processing, Tourism, and Urban Planning fields. In the physics field, Mingjing [71] measure and quantify the pH of soil by RF. 72 propose a model for extracting complex relationships between energy modulation and device efficiency. Zhang et al. [73] propose a model to accurately and effectively predict the UCS of LWSCC by a beetle antennae search algorithm for tuning the hyper-parameters of RF. The prediction of geotechnical parameters by typical RF is made by [74]. Creep index prediction by the RF algorithm to determine the optimal combination of variables is the main goal of [75]. In the text processing field, the comparison between RF and other

classifiers is presented by [76] for finding the best classifiers in the subject literature of text classification. The random forest is used as one of the base learners of the ensemble model for fake news detection by [77]. Analyzing the reviewer's comment for sentiment analysis is the main goal of [78]. Zhang et al. [79] propose two novel label flipping attacks to evaluate the robustness of NB under noise by random forest. Recognizing newspaper text by RF is done by [80]. Madichetty and Sridevi [81] use RF as one of the classifiers for detecting the damage assessment tweets. Madasu and Elango [82] use the typical RF for feature selection for sentiment analysis. Chang et al. [83] use online customer reviews for opinion mining by RF. Text classification by simple RF is the goal of [84]. Onan and Toçouglu [85] present a method for document clustering and topic modeling on massive open online courses. Sentiment analysis of technical words in English by the Gini index for feature selection is done by [86]. Beck [87] uses ensemble learning and deep learning for sentiment classification scheme with high predictive performance in massive open online courses' reviews. Onan [88] present a deep learning based approach to sentiment analysis. This approach uses TF-IDF weighted Glove word embedding with CNN LSTM architecture. Onan and Tocoglu [89] present an effective sarcasm identification framework on social media data by pursuing the paradigms of neural language models and deep neural networks. In the tourism field, Rodriguez-Pardo et al. [90] propose a method based on simple RF for predicting the behaviour of tourists. Predicting the travel time to reduce traffic congestion is the main goal of [91]. Jamatia et al.92 propose a method for tourist destinations' prediction. In urban planning, Baumeister et al.93 rank the urban forest characteristics for cultural ecosystem services supply by typical RF. Forecasting road traffic conditions in done by [94]. The simulation of urban space development by RF is presented by [95]. Investigating the information on a gross domestic product for the analysis of economic development is presented by [96]. Mei et al. [97] propose a method to identify the spatiotemporal commuting patterns of the transportation system. In this brief review, the mentioned references are categorized in terms of innovation and functionality.

As can be seen from Table 1, RF has a high range of applications and variations in scope. In contrast, both in

TABLE 1: Brief review of RF literature on functionality and innovation.

Type	Field	Paper
Functionality	Astronomy	[7], [8]
	Bioinformatics	[9], [10]
	Economics	[11], [12], [13]
	Global problem	[26], [27], [28]
	Healthcare	[32], [33], [34], [35], [36], [41], [98], [37], [39], [40], [42], [43], [45], [46], [47], [48], [49], [50], [51],
	Industrial	[52], [53], [54], [55], [56], [57], [58], [59], [60], [61], [62]
	Network	[63], [67], [68], [69], [99], [100]
	Physics	[71], [72]
	Text processing	[76], [77], [78], 80 [81], [82], [83], [84]
	Tourism	[91], [92]
	Urban planning	[93], [94], [95], [96], [97]
Innovative method	Economics	[14], [15]
	General	[16], [17], [18], [19], [101], [21], [22], [23], [24], [25]
	Global problem	[30], [31]
	Healthcare	[44]
	Industrial	[65]
	Network	[64]
	Physics	[73], [75]
	Text processing	[79], [86]

terms of quantity and quality, their innovations are often limited to set various parameters, and there is no significant innovation in the base learner combinations.

3. Learning Automata

Learning Automata (LA) is one of the learning algorithms that, after selecting different actions at different times, identify the best practices in terms of responses received from a random environment. LA selects an action from the set of actions in the vector of probabilities, and this action is evaluated in the environment. By using the received signal from the environment, the LA updates the probability vector and, by repeating this process, the optimal action is gradually identified. The classification problem can be formulated as a team of LA that operates collectively to optimize an objective function [102]. In Figure 1, the interaction of the learning automata and the environment is shown.

Finding the global optimum in the solution space is another advantage of using the LA. The LA can be formally represented by the quadruple

$$LA = \{\alpha, \beta, P, T\}, \tag{8}$$

in which

$$\alpha = \{\alpha_1, \alpha_2, \ldots, \alpha_r\} \tag{9}$$

is the set of actions (outputs) of the LA; in other words, the set of inputs of the environment.

$$\beta = \{\beta_1, \beta_2, \ldots, \beta_r\}, \tag{10}$$

is the set of inputs of the LA; in other words, the set of outputs of the environment.

$$P = \{p_1, p_2, \ldots, p_r\}, \tag{11}$$
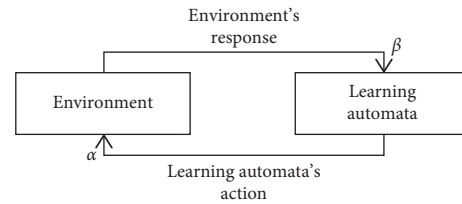
FIGURE 1: Interaction of learning automata with the environment.

is the probability vector of the LA actions and

$$P(n+1) = T[P(n), \alpha(n), \beta(n)], \tag{12}$$

is the learning algorithm.

In LA, three different models can be defined in the environment. In the P-Model, the environment presents the values of 0 or 1 as the output. In the Q-Model, the output values of the environment are discrete numbers between 0 and 1. In the S-Model, the output of the environment is the continuous value between 0 and 1. The selected actions by the LA are updated by both the signal received from the environment and using reward and penalty functions. The amount of allocated reward and penalty to the LA action can be defined in four ways: LRP, where the number of rewards and penalties are considered the same; LRεP in which the amount of penalty is several times smaller than the reward; LRI in which the penalty amount is considered 0; and LIP, where the reward amount is considered 0 [103].

At each instant n, the action probability vector pi(n) is updated by the linear learning algorithm given in equation (13) if the chosen action ai(k) is rewarded by the environment, and it is updated according to equation (14) if the chosen action is penalized [104].

$$\begin{cases} p_i(n+1) = p_i(n) + a[1 - p_i(n)], \\ p_j(n+1) = (1-a)p_j(n), \quad \forall j, \, j \neq i, \end{cases} \tag{13}$$

$$\begin{cases} p_j(n+1) = (1-b)p_i(n), \\ p_j(n+1) = \dfrac{b}{r-1} + (1-b)p_j(n), \quad \forall j; \, j \neq i, \end{cases} \tag{14}$$

, where "a" is the reward parameter, "b" is the penalty parameter, and "r" is the number of actions. The authors applied the LA in the proposed method, because:

(i) The LA presents an acceptable performance in uncertain situations.

(ii) The LA does search action in the probability space.

(iii) The LA requires simple feedback from the environment to optimize its state.

(iv) Since the LA has a simple structure, it has a simpler implementation in both software and hardware.

(v) The LA is not constrained to use accuracy criteria for optimization usage.

(vi) The LA is applicable in real-time usage since the LA is not involved with light computational complexity [105].

4. Proposed Method

The random forest is one of the methods of ensemble learning that all constructor classifiers are same type (i.e., decision tree). Therefore, the random forest is a homogeneous ensemble learning method. In this article, we intend to use the idea of reinforcement learning to increase the efficiency of random forest and add the ability to adapt to the conditions of the problem for this data mining algorithm. The details of the proposed method are described below.

The method proposed in this paper is based on the idea of reinforcement learning, and it employs the learning automata to implement the idea. The learning automata is the core of the proposed method, and by receiving feedback from the environment for each action, it updates the probability selection of the actions. In the proposed method, each base learner, all of which are decision tree, are considered as learning automata actions.

In the proposed method, the training data are first randomly divided into N sections. In this division, N corresponds to the number of trees we want to have in the forest. Unlike the random forest, in which the predictive model works by averaging or voting between trees, in the proposed method, the predictive model is created using learning automata, which forms the core of the algorithm. The block diagram of the proposed method is shown in Figure 2.

The preprocessing step in the proposed method is a general step, and based on what type of data the processing area is dealing with, the details of this phase are determined. In the proposed method, at first, similar to the random forest method, the training data are divided into the number of

base learners and randomly injected into the base learners. The difference between this step and the similar step in the random forest is that all the features in the feature vector are given to all base learners, and the feature replacement option is removed.

After the first run, the prediction models are created in the base learners and placed in a pool that is actually an interactive environment with the learning automata. The results obtained from the base learners for each new sample are given in the form of a reinforcement signal to the learning automata, which we know as the primary feedback of the environment. Depending on whether the received reinforcement signal is a reward or a penalty, the chances of selecting each of the base learners, -which they are the actions of the learning automata - are updated. It should be noted that the initial probability of selecting these actions is considered equal at the start. If we have R base learners to form the ensemble, the probability of the initial selection of each of them is equal to

$$p(DT_r) = (1/R). \tag{15}$$

It is clear that the sum of the probabilities of all actions will be equal to 1.

$$\sum_{i=1}^{R} (pDT) = 1. \tag{16}$$

The initial probability of selecting actions is considered equal because all of them are homogeneous in terms of separating power.

In the proposed method, integration of the base learners is performed by the LA. Therefore, for each input in the test set, a linear LA is defined, and the action of each LA corresponds to selecting the base learners. The process of running base learners and receiving feedback from the environment continues until the probability of selecting actions converges to one of the base learners, or the number of repetitions for learning automata exceeds the predetermined limit. Once the probability of selections converges, then the result of the base learner, which has the highest probability of selection, is determined as the result of the ensemble for that particular input. In such a case, finding the global optimal is guaranteed by the algorithm, and because all the features in the feature vector are examined, rare modes are also covered, and the ability to adapt to the conditions of the problem space and independence from the domain is stabilized. In the proposed method, the random selection of subsets causes interdependence between trees. The depth of all the decision trees in the proposed method is considered equal. Each decision tree divides the training data differently at the leaf level. The pseudocode of the proposed method is shown in Algorithm 2.

In the learning automata block in Figure 2, there are two functions called the reward function and penalty function. Activation of one of these two functions is based on the type of reinforcement signal received from the environment. The received signal from the environment determines whether the result of the base learner activity or the selected action in the previous step was useful or not. If the result is useful, that
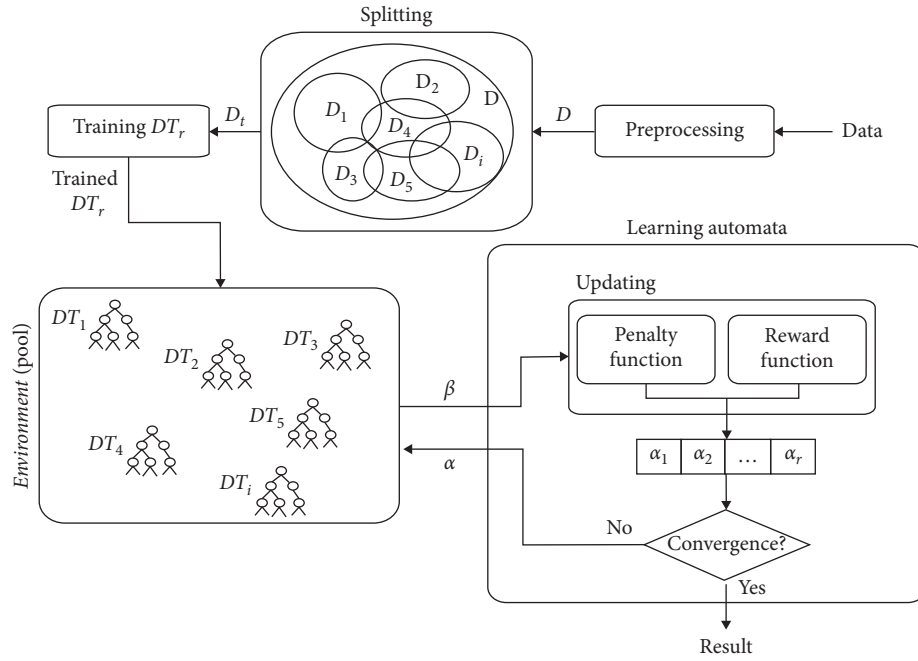
FIGURE 2: The block diagram of the proposed method.

Input D = {(**x₁, y₁**), (**x₂, y₂**), . . ., (**x_N, y_N**),} denote the training data with $x_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,p})$

(1) **Output** classified test data
(2) **Assumption**
(3) 　**LA** : Learning automata
(4) 　**DT_r** = {DT₁, DT₂, . . . , DT_R} denote the base learners
(5) 　**α_i**: LA action//Choose DT_r
(6) 　**a**: Reward parameter
(7) 　**b**: Penalty parameter
(8) 　**Pool** : All the trained base learners
(9) **Algorithm**
(10) 　**For** r = 1 to R **do**
(11) 　　Create a dataset **D_t**, by sampling (**N/R**) items, randomly with replacement from **D**
(12) 　　Train **DT_r** using **D_t**, and add to the **pool**
(13) 　**end**//for
(14) 　**For** each test sample
(15) 　　{
(16) 　　**LA** = new **LA**//Create an LA object from LA class
(17) 　　**While** ((**LA** convergences to an action) or (**LA** exceeds predefined iteration number))
(18) 　　{
(19) 　　Select one of the actions at random and execute it, by the **LA**, Let it be **α_i**
(20) 　　**If** (**α_i** predicts the new test sample correctly) **then**//Update the probability of selection vector
(21)
$$\begin{cases} p_i(n+1) = p_i(n) + a[1 - p_i(n)] \\ p_j(n+1) = (1-a)p_j(n), \quad \forall j, j \neq i \end{cases} //\text{reward the selected } \pmb{\alpha_i}$$
(22)
　　　else
(23)
$$\begin{cases} p_j(n+1) = (1-b)p_i(n), \\ p_j(n+1) = (b/R - 1) + (1-b)p_j(n), \quad \forall j, j \neq i, \end{cases} //\text{Penalty the selected } \pmb{\alpha_i}$$
(24)
　　　}//end while
(25) 　　}//end for
(26) **Return DT_r**
(27) Classified test data = the prediction of **DT_r**
(28) **End.**//algorithm

ALGORITHM 2: The pseudocode of the proposed method.

action must be rewarded or, in other words, increase the probability of its selection. The increase in the probability of the selected action is determined by the parameters "a" and "b," which are called the reward parameter and the penalty parameter, respectively.

To comply with (16), that is, the sum of the probabilities of all actions being equal to one, the probability of all other actions is reduced according to the size of the parameter "a." If the result of the selected action is not useful, that action must also be penalized. In other words, the probability of that action must be reduced. To do this, the probability of selecting that action is reduced to the size of parameter "b," and as a rewarding mode, and to observe (16), the probability of selecting other actions is increased by the size of the parameter "b."

In the proposed method, the learning automata model environment is assumed to be the P-Model, where the environment defines zero and one values as outputs. Zero means reward, and one means penalty. If the correct answer is received from the selected base learner by the LA, the action of choice will be rewarded; otherwise, it will be penalized.

5. Evaluation

In order to thoroughly evaluate the efficiency of the proposed method, in this section, the details of the evaluation of the proposed method are presented separately from the data used and the experimental results.

5.1. Datasets. In order to evaluate the proposed method and to create an environment with the dynamic behaviour of data, different domains of applications have been selected. As mentioned in the previous sections, dynamic behaviour refers to the different results that an instance exhibits in different environmental conditions. Variety in the results of different environments is created by a specific domain. Text data are one of the most well-known types of data that exhibit such dynamic behaviour. In other words, these types of data are one of the optimal options for creating a dynamic environment, which proves the adaptability of the proposed method. The details of the selected data for the evaluation phase are shown in Table 2.

5.2. Experimental Result. In order to evaluate the proposed method, eighteen datasets in different domains introduced in the previous section have been used. In the literature on learning automata, different modes have been considered for tuning learning automata; in this paper, three modes have been used to evaluate the proposed method. The LIP mode is not considered due to poor results. The evaluation results of each of the LRI, LRεP, and LRP modes are shown in separate figures. In order to determine the optimal value for the reward and penalty parameters, six text datasets have been selected. The reason for this choice is the high diversity in the behaviour of textual data as well as a large number of samples and a large number of features of these six datasets. In the LRI mode, the value of the penalty parameter is

considered to be zero, and the results of the proposed method in this mode are shown in Figure 3.

Based on the literature on learning automata in the LRεP mode, the value of the penalty parameter is considered to be much smaller than the value of the reward parameter. The results of the proposed method are shown in the LRεP mode in Figure 4.

As mentioned in the learning automata section, in the LRP mode, the values of the penalty and reward parameters are considered equal. The results of the proposed method in this mode are also shown in Figure 5.

A comparison of the results obtained from the implementation of the proposed method in three adjustable modes for learning automata shows that the settings on the LRP mode have resulted in the highest accuracy for identification. Then there are LRεP and LRI modes. In the LRεP mode, the setting $a = 0.01$, $b = 0.01$ is not considered, because these values are equal to the first values set in the LRP mode, and in order to prevent duplication of results in different tables, these settings have been removed from the LRεP mode. For this reason, the number of experiments performed on LRεP mode evaluations is one less than the other two. Considering that the settings of reward and penalty parameters in the LRP mode with the values of $a = 0.5$, $b = 0.5$ have resulted in the highest efficiency, evaluation has been done on other datasets with these settings. A comparison of the proposed method and similar approaches in the subject literature is shown in Table 3.

As can be seen in Table 3 from the point of view of accuracy, the proposed method offers better performance than the methods available in the subject literature, which indicates an improvement in the aggregation model of the base learners. This improvement is due to the use of reinforcement learning ideas of the method of aggregation of basic classifiers, which is known as base learner. The use of reinforcement learning ideas has improved the ability of the created ensemble, and it improved the ability to address issues in which data exhibit dynamic behaviour. The results of experiments performed on different data confirm the capabilities added to the random forest by the proposed method. As mentioned earlier, in the field of opinion mining, the type of text data is the most obvious data that exhibit such dynamic behaviour. Therefore, the optimal values for the reward and penalty parameters have been determined in these types of data, and these settings have been used for other types of data.

In addition to the accuracy criterion, other statistical criteria have been examined to evaluate the proposed method. As can be seen in Table 4, the proposed method has shown better results in both positive and negative classes than the methods available in the literature. Among the statistical criteria, Precision (P) determines the exactness of the results obtained from the classifier, and Recall (R) determines the completeness of the results obtained from the classifier. The results obtained from the test in the mentioned statistical criteria show that the proposed method has a high performance.

TABLE 2: Details of textual data used for evaluation.

Domain	Name	# Feature	# Instance
Text	Stanford—Sentiment 140 corpus [106]	Bag of word	1600000
	Large dataset of movie reviews [107]	Bag of word	50000
	Sentence polarity dataset v1.0 [108]	Bag of word	10662
	Internet movie database [105]	Bag of word	1400
	Yelp review [105]	Bag of word	598000
	Amazon review [105]	Bag of word	1000000
Healthcare	Heart disease dataset [105]	13	200
	Breast cancer dataset [105]	30	569
	Arrhythmia dataset [105]	279	454
	Parkinson dataset [105]	45	241
	Caesarean section dataset [105]	5	81
	Gene expression dataset [105]	255	801
	Diabetes dataset [105]	7	765
	Statlog (heart) dataset [105]	13	271
Physical	Ionosphere dataset [105]	34	352
	Sonar, mines vs. rocks dataset [105]	60	208
Sound	Voice dataset [105]	20	3168
	Emotions from music dataset [105]	28	592

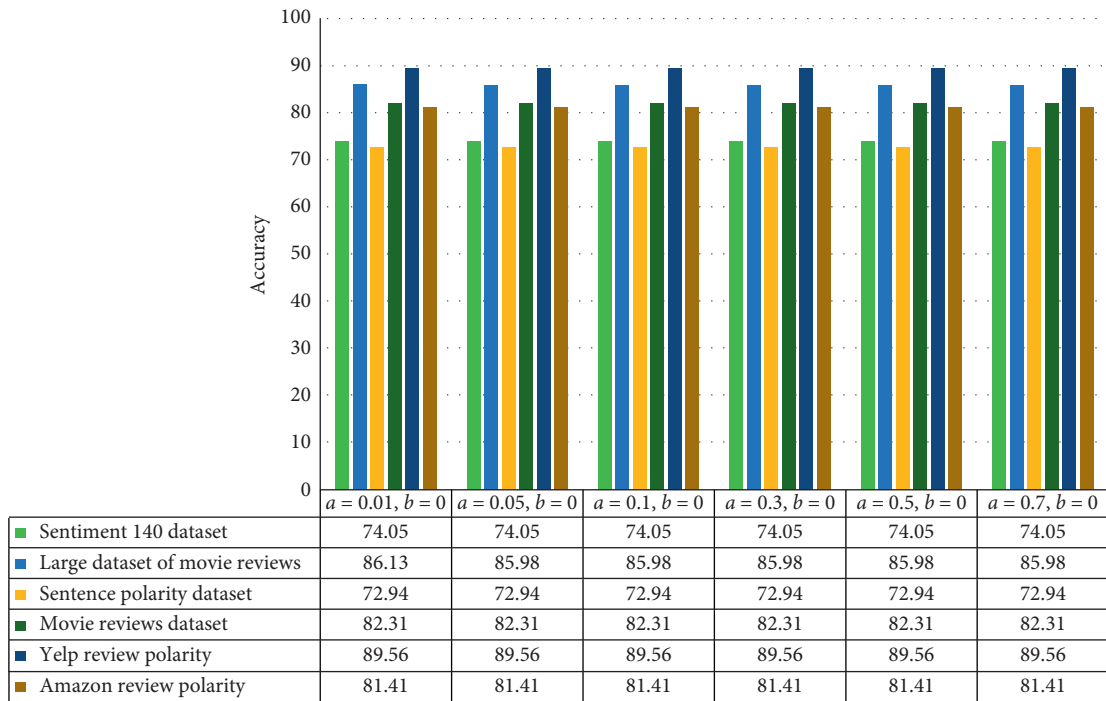|	$a = 0.01, b = 0$	$a = 0.05, b = 0$	$a = 0.1, b = 0$	$a = 0.3, b = 0$	$a = 0.5, b = 0$	$a = 0.7, b = 0$
■ Sentiment 140 dataset	74.05	74.05	74.05	74.05	74.05	74.05
■ Large dataset of movie reviews	86.13	85.98	85.98	85.98	85.98	85.98
■ Sentence polarity dataset	72.94	72.94	72.94	72.94	72.94	72.94
■ Movie reviews dataset	82.31	82.31	82.31	82.31	82.31	82.31
■ Yelp review polarity	89.56	89.56	89.56	89.56	89.56	89.56
■ Amazon review polarity	81.41	81.41	81.41	81.41	81.41	81.41

FIGURE 3: The results of the proposed method in LRI mode.

6. Discussion

In this section, more details of the proposed method are explained along with the reasons for the need to address these details. These include the details of the preprocessing step, tuning the learning automata parameters, as well as ranking the set of these parameters based on their performance.

6.1. Preprocessing. As explained in the proposed method section, the preprocessing step is a general step. In order for
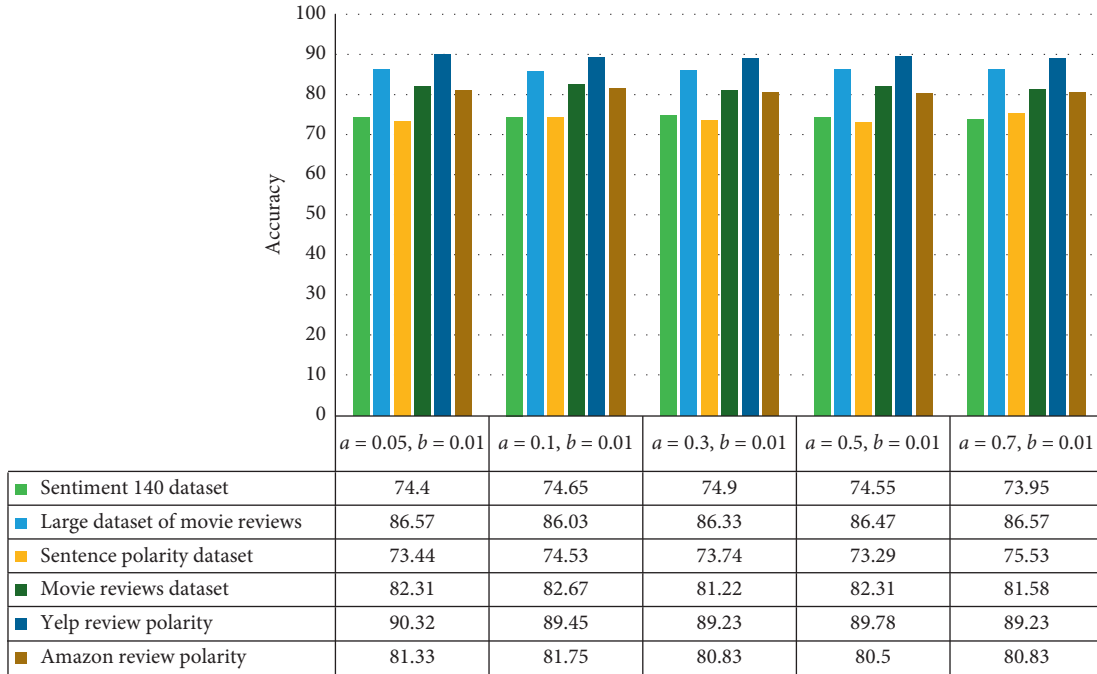
	$a = 0.05, b = 0.01$	$a = 0.1, b = 0.01$	$a = 0.3, b = 0.01$	$a = 0.5, b = 0.01$	$a = 0.7, b = 0.01$
■ Sentiment 140 dataset	74.4	74.65	74.9	74.55	73.95
■ Large dataset of movie reviews	86.57	86.03	86.33	86.47	86.57
■ Sentence polarity dataset	73.44	74.53	73.74	73.29	75.53
■ Movie reviews dataset	82.31	82.67	81.22	82.31	81.58
■ Yelp review polarity	90.32	89.45	89.23	89.78	89.23
■ Amazon review polarity	81.33	81.75	80.83	80.5	80.83

FIGURE 4: The results of the proposed method in the LRεP mode.

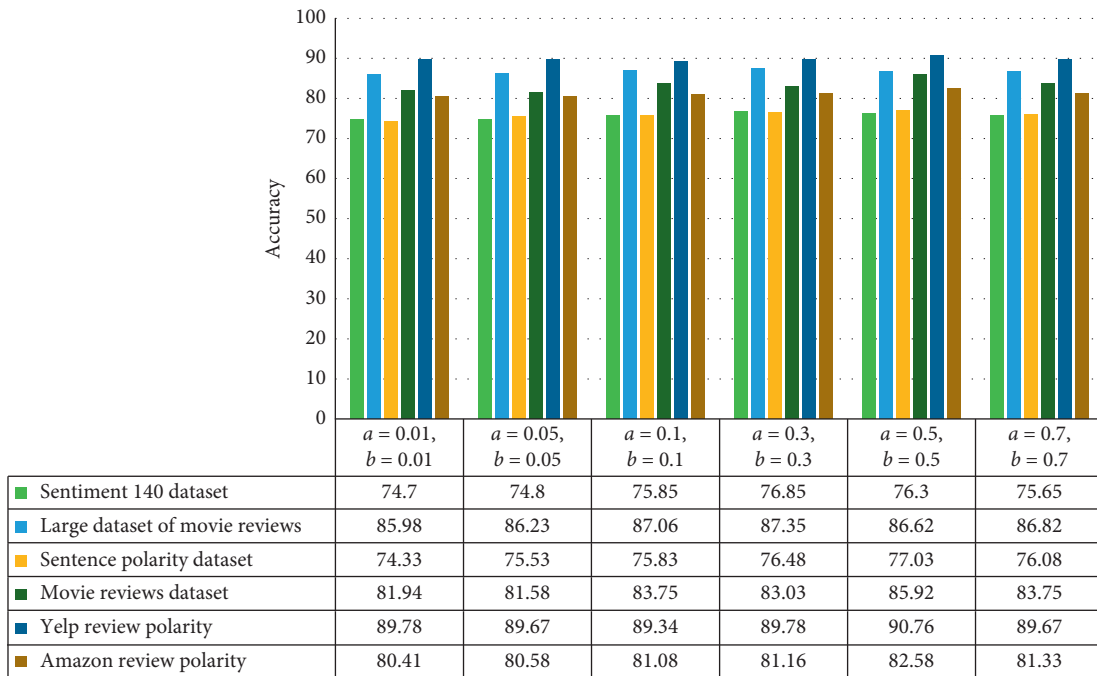|	$a = 0.01,$ $b = 0.01$	$a = 0.05,$ $b = 0.05$	$a = 0.1,$ $b = 0.1$	$a = 0.3,$ $b = 0.3$	$a = 0.5,$ $b = 0.5$	$a = 0.7,$ $b = 0.7$
■ Sentiment 140 dataset	74.7	74.8	75.85	76.85	76.3	75.65
■ Large dataset of movie reviews	85.98	86.23	87.06	87.35	86.62	86.82
■ Sentence polarity dataset	74.33	75.53	75.83	76.48	77.03	76.08
■ Movie reviews dataset	81.94	81.58	83.75	83.03	85.92	83.75
■ Yelp review polarity	89.78	89.67	89.34	89.78	90.76	89.67
■ Amazon review polarity	80.41	80.58	81.08	81.16	82.58	81.33

FIGURE 5: The results of the proposed method in the LRP mode.

the evaluation, different data from different domains were examined. The preprocessing of textual data, along with the relevant details, is described below. It should be noted that preprocessing for other types of data, such as feature extraction, feature selection, normalization, noise removal, and other related preprocessing, has not been performed because all of them are taken as clean data from the UCI Repository [109]. And their basis for accuracy is based on previous research works that have used these data.

In order to prepare textual data for the main process, the opinion mining domain is selected and the related preprocessing is as follows. The details of the preprocessing step for text data in opinion mining are shown in Figure 6.

TABLE 3: Comparison of the proposed method with similar approaches in the subject literature.

	Dataset	Averaging	Majority Voting	Random Forest	Our Method
Text	Sentiment140 dataset	74.54	75.50	74.30	**76.30**
	Large dataset of movie reviews	86.28	86.86	86.42	**86.62**
	Sentence polarity dataset	73.75	74.63	73.38	**77.03**
	Movie reviews dataset	81.58	81.58	81.67	**85.92**
	Yelp review polarity	89.47	90.32	89.74	**90.76**
	Amazon review polarity	80.86	81.66	80.97	**82.58**
Healthcare	Heart disease dataset	58.00	57.50	57.50	**65.00**
	Breast cancer dataset	97.41	97.36	96.49	**98.24**
	Arrhythmia dataset	80.71	85.71	81.31	**85.71**
	Parkinson dataset	63.95	64.58	64.58	**68.75**
	Caesarean section dataset	60.31	62.50	43.75	**68.75**
	Gene expression dataset	95.59	95.62	96.27	**98.75**
	Diabetes dataset	75.77	75.32	74.67	**76.62**
Physical	Statlog (heart) data set	81.20	81.48	79.62	**85.18**
	Ionosphere dataset	91.05	91.54	92.95	**95.77**
	Sonar, mines vs. rocks dataset	85.23	85.71	73.80	**88.09**
Sound	Voice dataset	76.38	76.18	76.49	**88.95**
	Emotions from music dataset	78.23	78.15	82.35	**84.03**

TABLE 4: Comparison of statistical criteria.

Method	Positive class			Negative class			Method	Positive class			Negative class		
	P (%)	R (%)	F1 (%)	P (%)	R (%)	F1 (%)		P (%)	R (%)	F1 (%)	P (%)	R (%)	F1 (%)
Sentiment140 dataset							*Parkinson dataset*						
MV	72.44	70.35	71.38	70.90	72.96	71.92	MV	69.23	66.67	67.92	**59.09**	61.9	60.47
RF	72.44	74.58	73.49	76.46	74.43	75.43	RF	69.23	66.67	67.92	**59.09**	61.9	60.47
OM	**75.20**	**80.36**	**77.69**	**81.17**	**76.16**	**78.59**	OM	**76.92**	**68.97**	**72.73**	59.09	68.42	63.41
Large dataset of movie reviews							**Caesarean section data set**						
MV	87.50	87.33	87.41	87.41	87.59	87.50	MV	55.56	71.43	62.5	**71.43**	55.56	62.5
RF	85.83	76.16	80.70	73.37	83.93	78.29	RF	33.33	50	40	57.14	40	47.06
OM	**87.80**	**87.97**	**87.88**	**88.10**	**87.93**	**88.01**	OM	**66.67**	**75**	**70.59**	**71.43**	**62.5**	**66.67**
Sentence polarity dataset							**Gene expression dataset**						
MV	75.63	72.83	74.20	72.80	75.61	74.18	MV	92.31	94.12	93.2	97.25	96.36	96.8
RF	74.62	67.78	71.08	65.95	72.94	69.27	RF	94.23	93.23	94.23	97.25	97.25	97.25
OM	**76.04**	**73.29**	**74.64**	**73.29**	**76.04**	**76.64**	OM	**98.08**	**98.08**	**98.08**	**99.08**	**99.08**	**99.08**
Movie reviews dataset							**Diabetes dataset**						
MV	83.10	82.52	82.81	81.84	82.09	81.78	MV	90.29	76.86	83.04	45.1	69.7	54.76
RF	77.46	71.43	74.32	76.41	73.98	70.54	RF	90.29	76.23	82.67	43.14	68.75	53.01
OM	**84.51**	**83.33**	**83.92**	**82.22**	**83.46**	**82.84**	OM	**91.26**	**77.69**	**83.93**	**47.07**	**72.73**	**57.14**
Yelp review polarity							**Voice dataset**						
MV	**90.85**	88.04	88.42	87.11	**90.11**	88.59	MV	64.58	76.09	69.86	84.85	76.24	80.31
RF	80.64	85.94	83.21	86.22	81.00	83.53	RF	64.58	76.75	70.14	85.4	76.35	80.62
OM	90.00	**89.43**	**89.71**	88.89	89.49	**89.19**	OM	**87.82**	86.55	**87.18**	89.81	90.81	90.3
Amazon review polarity							**Emotions from music dataset**						
MV	82.68	79.45	81.03	**79.38**	82.62	80.97	MV	83.58	78.87	81.16	71.15	77.08	74
RF	79.12	74.56	76.77	73.98	87.61	76.22	RF	88.06	**81.94**	84.89	**75**	82.98	78.79
OM	**83.70**	**79.52**	**81.56**	79.21	**83.45**	**81.28**	OM	**92.54**	81.58	**86.71**	73.08	**88.37**	**80**
Heart disease dataset							**Sonar, mines vs. Rocks dataset**						
MV	**61.90**	59.09	60.47	52.63	55.56	54.05	MV	**87.5**	**87.5**	**87.5**	**83.33**	**83.33**	**83.33**
RF	**61.90**	59.09	60.47	52.63	55.56	54.05	RF	79.17	76	77.55	66.67	70.59	68.57
OM	**61.90**	**68.42**	**65.00**	**68.42**	**61.90**	**65.00**	OM	**87.5**	**87.5**	**87.5**	**83.33**	**83.33**	**83.33**
Breast cancer data set							**Statlog (heart) data set**						
MV	95.74	97.83	96.77	**98.51**	97.06	97.78	MV	**79.19**	79.19	79.19	83.33	83.33	83.33
RF	**97.87**	93.88	95.83	95.52	98.46	96.97	RF	75	78.26	76.6	83.33	80.65	81.67
OM	**97.87**	**97.87**	**97.87**	**98.51**	**98.51**	**98.51**	OM	**79.19**	**86.36**	**82.61**	**90**	**84.38**	**87.1**
Arrhythmia data set							**Ionosphere data set**						
MV	88.37	**82.61**	85.39	83.33	88.89	86.02	MV	82.14	95.83	88.46	97.67	89.36	93.33
RF	83.72	78.26	80.9	79.17	79.74	73.85	RF	85.71	96	90.57	97.67	91.3	94.38
OM	**93.02**	80	**86.02**	79.19	**92.68**	**85.39**	OM	**89.29**	**100**	**94.34**	**100**	**93.48**	**96.63**

P, *R*, and F1 refer to Precision, Recall, and F1-score. MV: majority voting, RF: random forest, and OM: our method.
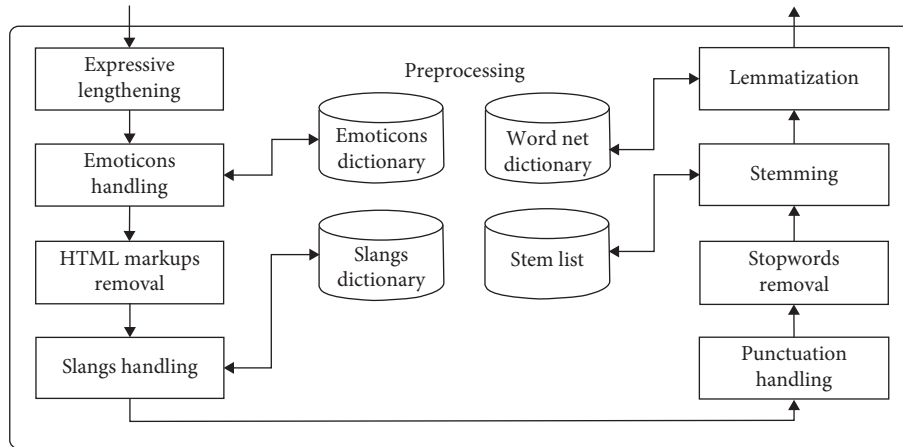
FIGURE 6: Details of the preprocessing step for text data.

Expressive Lengthening. Word lengthening or word stretching refers to the words that are elongated to express a particular emotion strongly, and the words with wrong spellings are corrected and replaced with their original words.

Emoticons Handling. It refers to the emoticons mentioned in the text that are replaced with their meaning, which makes it easier to analyze the emoticons.

HTML Markups Removal. HTML markups presented in the text are removed as they do not have any sentimental value attached to it.

Slangs Handling. The slangs are used for writing a given word, in short syllables, which depict the same meaning but save the time of typing. In slangs handling, the slangs presented in the text are replaced with their original words.

Punctuation Handling. Punctuations are used in a text to separate sentences and their elements, and to clarify their meaning. At punctuation handling, once the apostrophes are handled, all the remaining punctuations and numbers are removed.

Stopwords Removal. Stopwords do not carry much meaning and have no importance in the text. Stopwords are removed to get a simplified text.

Stemming. It refers to finding out the root or stem of a word. Removing various suffixes to reduce the number of words is the purpose of stemming.

Lemmatization. It returns the base or dictionary form of a word, which is known as the lemma. It is very similar to stemming, but it is more akin to synonym replacement.

BoW creation. The bag of word creation is the latest preprocess that is performed on the text preparation.

6.2. Tuning the Parameters of Reward and Penalty. In the subject literature of the learning automata, three different modes have been defined to tune the parameters of reward and penalty. In the proposed method, in which the idea of

reinforcement learning is implemented using learning automata, all three adjustable modes of the parameters of reward and penalty are examined. The results of these three modes were presented in the experimental result section. In this paper, Friedman test statistical verification is used to determine which mode and which settings are best adjustable for the reward and penalty parameters. The values set for parameters "*a*" and "*b*" are shown in Table 5. Determining the numerical value of these parameters is based on the subject literature of learning automata. Of course, a wide variety of values can be considered for these two parameters. In this paper, an attempt has been made to tune the parameters in such a way that all the modes are considered so that they can be used to prove the efficiency of the proposed method compared to the previous methods.

6.3. Ranking. Friedman test statistical verification [110] is a ranking method that, the difference between the ranks assigned to each of the input samples, determines the optimal level of each option. In this paper, this verification method has been used to determine the optimal value of reward and penalty parameters as well as to compare the proposed method with the conventional methods in the subject literature of ensemble learning. The results are shown in Table 6.

As can be seen in Table 6, there is a significant difference between the rankings of the proposed method and the rankings of the traditional methods, which indicate an improvement in the efficiency of the proposed method compared to other methods. Among the three modes considered for tuning reward and penalty parameters, it is observed that the rankings have increased in LRI, LReP, and LRP modes, respectively. In the LRP mode, where the values of the reward and penalty parameters are considered the same, the highest efficiency is also observed. There is a significant difference between the Mean Rank of the best set of the reward and penalty parameters in the proposed method and this rank in the random forest method. The difference between the ranks is proof that the proposed method is optimal versus the traditional methods of

TABLE 5: Numerical values tuned for reward and penalty parameters.

Mode		Parameter					
LRI	a	0	0.1	0.1	0.3	0.5	0.7
	b	0	0	0	0	0	0
LRεP	a	0.1	0.1	0.3	0.5	0.7	
	b	0	0	0	0	0	
LRP	a	0	0.1	0.1	0.3	0.5	0.7
	b	0	0.1	0.1	0.3	0.5	0.7

TABLE 6: Friedman test statistical verification results for ranking the parameters of reward and penalty and comparing the proposed method with the literature.

Method	Tuning	Mean rank	Final rank
LRP	$a = 0.5$, $b = 0.5$	19.17	1
LRP	$a = 0.3$, $b = 0.3$	16.83	2
LRP	$a = 0.7$, $b = 0.7$	15.58	3
MV	Majority voting	14.67	4
LRP	$a = 0.1$, $b = 0.1$	13.92	5
LReP	$a = 0.05$, $b = 0.01$	12.17	6
LReP	$a = 0.1$, $b = 0.01$	11.83	7
LReP	$a = 0.5$, $b = 0.01$	10.08	8
LRP	$a = 0.05$, $b = 0.05$	9.58	9
RF	Random forest	9.17	10
LRP	$a = 0.01$, $b = 0.01$	8.75	11
LIR	$a = 0.01$, $b = 0$	8.42	12
LIR	$a = 0.05$, $b = 0$	7.67	13
LIR	$a = 0.1$, $b = 0$	7.67	13
LIR	$a = 0.3$, $b = 0$	7.67	13
LIR	$a = 0.5$, $b = 0$	7.67	13
LIR	$a = 0.7$, $b = 0$	7.67	13
AV	Averaging	7.58	14
LReP	$a = 0.3$, $b = 0.01$	7.17	15
LReP	$a = 0.7$, $b = 0.01$	6.75	16

(a)

(b)
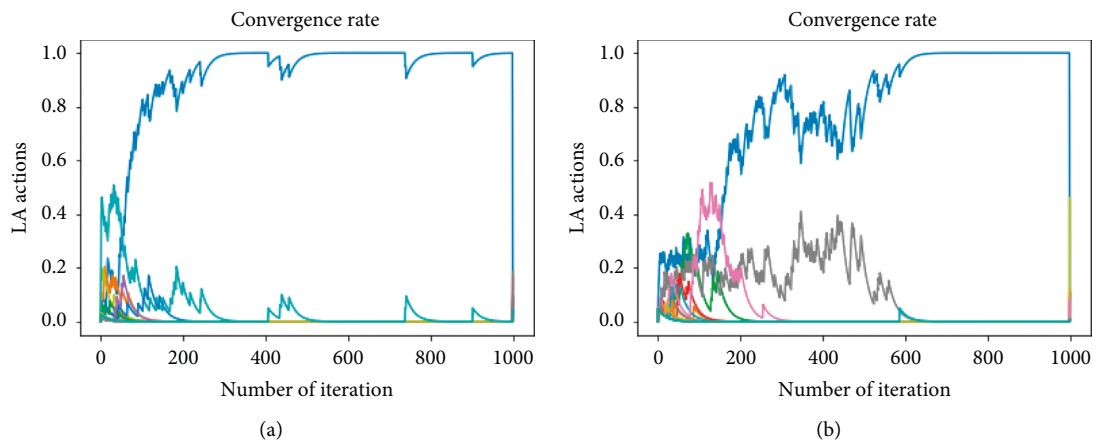
FIGURE 7: Continued.

(c)

(d)

(e)

(f)

(g)
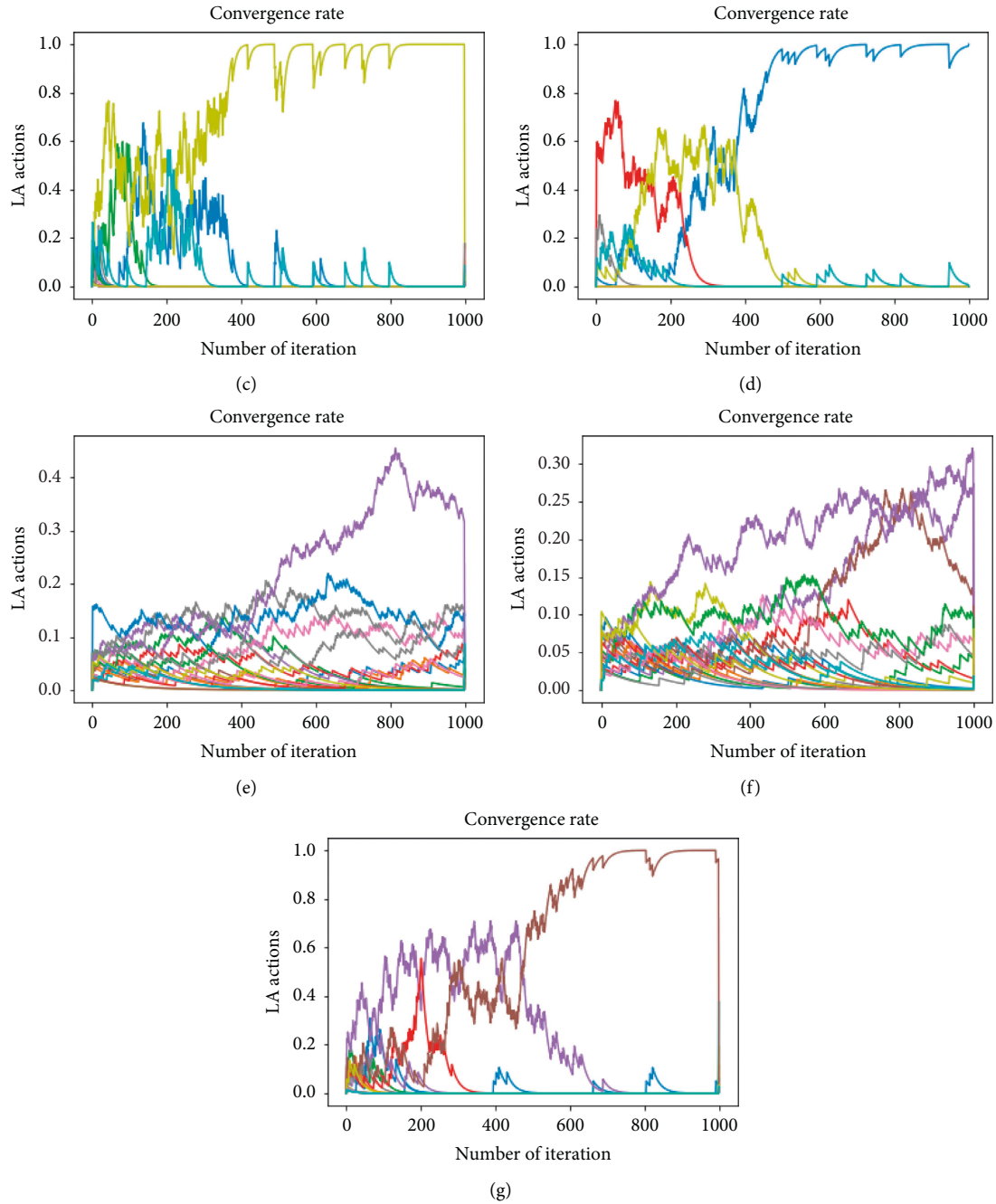
FIGURE 7: Convergence rate for different reward and penalty parameters. (a) $a = 0.5$, $b = 0.5$; (b) $a = 0.3$, $b = 0.3$; (c) a = 0.7, b = 0; (d) $a = 0.1$, $b = 0.1$; (e) $a = 0.01$, $b = 0$; (f) $a = 0.05$, $b = 0.05$; (g) $a = 0.3$, $b = 0$.

aggregating classifiers to achieve a strong classification method.

6.4. Checking Convergence Rate. To more accurately address the proposed method in terms of efficiency, LA convergence has been investigated. Figure 7 shows the convergence of LA actions for different amounts of reward and penalty variables. In most of the different settings for these two parameters, the convergence rate is high, and convergence to one of the actions usually occurs before reaching a certain

number of iterations. As shown in Table 5, convergence at a lower rate occurred in some of the other settings that scored lower on the Friedman test.

6.5. Noise Resistance. In order to more accurately evaluate the proposed method and determine the resistance of the proposed method to noise, another evaluation has been performed on the data presented in the previous section. This evaluation was performed by injecting 20% noise into clean data. The results of the evaluation on noisy data show
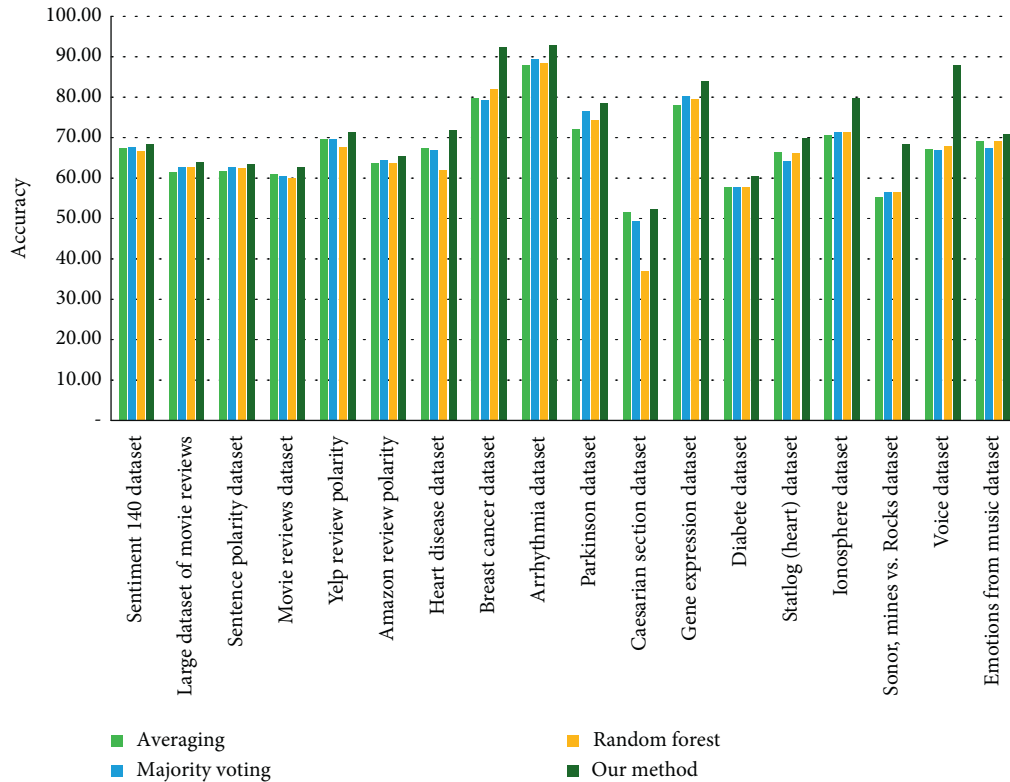
FIGURE 8: The evaluation of the proposed method in the presence of noise.

that the proposed method, due to the use of learning automata, has high adaptability to the problem conditions, and in the presence of noise, contrary to conventional methods in the literature, the proposed method does not suffer a sharp decline, and in such conditions, it shows high efficiency compared to traditional methods. The evaluation of the proposed method in the presence of noise is shown in Figure 8.

7. Conclusion and Future Work

Base learner aggregation in ensemble learning should be done in such a way that the following points are met. First point: selecting a base learner leads to the highest performance achievable in the current situation. Second point: if the situation changes due to the dynamics of the problem, the structure of the ensemble will change in such a way that it has the greatest amount of compatibility with the conditions of the new environment. Therefore, in order to meet the above points and achieve an ensemble that is able to adapt to the dynamic conditions of the problem, in this paper, a new method based on the idea of reinforcement learning is proposed to integrate the base learners in the random forest. In the proposed method, learning automata is used to receive feedback from the environment and perform actions on it. The general procedure is to receive feedback from the environment, where the environment is a set of base learners that we intend to combine to achieve a better performance than individual base learners. Learning automata actions include choosing one of the base learners as the best base

learner. The choice of action is based on receiving feedback from the environment. This causes the dynamic behaviour of data to be covered by using the idea of reinforcement learning. On the other hand, given that at each stage, learning automata strives to achieve the highest amount of achievable rewards, it is guaranteed to find the global optima in the problem space. Adaptability is another advantage of the proposed method compared to similar methods in the subject literature.

Due to the fact that in each step learning automata operates based on environmental conditions and received feedback from the environment, the ability to adapt to the problem is met. The results of the evaluations performed in different data show that the proposed method has the ability to achieve all the desired items mentioned above. Despite the fact that, unlike the random forest mechanism, all features are injected into all base learners in the proposed method, the efficiency of the proposed method in dealing with large-volume data has not decreased, and the results are more favorable than the classical methods. The proposed method is independent of the data type and has the ability to handle any other type of data in any field. In order to substantiate this claim, and in order to evaluate the proposed method, different types of data have been chosen. However, there are no restrictions on the proposed method for dealing with different types of data. In this paper, a new method for aggregating the base learners of the random forest using learning automata is proposed. Determining the optimal value for the parameters of reward and penalty in the form of self-tuning is one of the future works that the authors intend to do.

Data Availability

The authors declare that all the data are available publicly at the UCI repository.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] C. Zhang and Y. Ma, *Ensemble Machine Learning: Methods and Applications*, Springer, New York, NY, USA, 2012.

[2] A. Galicia, R. Talavera-Llames, A. Troncoso, I. Koprinska, and F. Martínez-Álvarez, "Multi-step forecasting for big data time series based on ensemble learning," *Knowledge-Based Systems*, vol. 163, pp. 830–841, 2019.

[3] M. Savargiv, B. Masoumi, and M. R. Keyvanpour, "A new ensemble learning method based on learning automata," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–16, 2020.

[4] M. Kang, J. Ahn, and K. Lee, "Opinion mining using ensemble text hidden Markov models for text classification," *Expert Systems with Applications*, vol. 94, pp. 218–227, 2018.

[5] Y. Zhang, D. Miao, J. Wang, and Z. Zhang, "A cost-sensitive three-way combination technique for ensemble learning in sentiment classification," *International Journal of Approximate Reasoning*, vol. 105, pp. 85–97, 2019.

[6] T.-H. Lee, A. Ullah, and R. Wang, *"Bootstrap Aggregating and Random Forest*, pp. 389–429, Springer, New York, NY, USA, 2020.

[7] J. Markel and A. J. Bayless, "Performance of random forest machine learning algorithms in binary supernovae classification," 2019, http://arxiv.org/abs/1907.00088.

[8] B. ChenZ. Li et al., "Forest signal detection for photon counting LiDAR using Random Forest," *Remote Sensing Letters*, vol. 11, no. 1, pp. 37–46, 2020.

[9] W. Pang, X. Liu, Z. Wang, Y. Fan, and J. Wang, "Predicting RNA molecular specific hybridization via random forest," in *Proceedings of the 2019 IEEE 7th International Conference on Bioinformatics and Computational Biology (ICBCB)*, pp. 35–38, Hangzhou, China, 2019.

[10] M. F. Darmawan, A. F. Zainal Abidin, S. Kasim, T. Sutikno, and R. Budiarto, "Random forest age estimation model based on length of left hand bone for Asian population," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 1, p. 549, 2020.

[11] S. Park, J. Moon, S. Jung, S. Rho, S. W. Baik, and E. Hwang, "A two-stage industrial load forecasting scheme for day-ahead combined cooling, heating and power scheduling," *Energies*, vol. 13, no. 2, p. 443, 2020.

[12] J. Dai, T. Wang, and S. Wang, "A deep forest method for classifying e-commerce products by using title information," in *Proceedings of the 2020 International Conference On Computing, Networking And Communications (ICNC)*, pp. 1–5, Big Island, HI, USA, 2020.

[13] M. Papoušková and P. Hajek, "Modelling loss given default in peer-to-peer lending using random forests," in *Proceedings of the Intelligent Decision Technologies 2019*, pp. 133–141, Springer, Malta, Europe, 2020.

[14] D. Borup, B. J. Christensen, N. Mühlbach, and M. S. Nielsen, *The Effects Of Targeting Predictors In A Random Forest Regression Model*, 2020.

[15] S. Sikdar, V. Kadiyali, and G. Hooker, *Price Dynamics on Amazon Marketplace: A Multivariate Random Forest Variable Selection Approach*, 2019.

[16] L. Giffon, C. Lamothe, L. Bouscarrat, P. Milanesi, F. Cherfaoui, and S. Koço, *Pruning Random Forest with Orthogonal Matching Trees*, 2020.

[17] J. L. Speiser, B. J. Wolf, D. Chung, C. J. Karvellas, D. G. Koch, and V. L. Durkalski, "BiMM forest: a random forest method for modeling clustered and longitudinal binary outcomes," *Chemometrics and Intelligent Laboratory Systems*, vol. 185, pp. 122–134, 2019.

[18] N. Mohapatra, K. Shreya, and A. Chinmay, "Optimization of the random forest algorithm," in *Advances In Data Science And Management*, pp. 201–208, Springer, New York, NY, USA, 2020.

[19] Q. Ji, T. Zhu, and D. Bao, "A hybrid model of convolutional neural networks and deep regression forests for crowd counting," *Applied Intelligence*, pp. 1–15, 2020.

[20] B. Santra, A. Paul, and D. P. Mukherjee, "Deterministic dropout for deep neural networks using composite random forest," *Pattern Recognition Letters*, vol. 131, pp. 205–212, 2020.

[21] M. A. Ganaie, M. Tanveer, and P. N. Suganthan, "Oblique decision tree ensemble via twin bounded SVM," *Expert Systems with Applications*, vol. 143, Article ID 113072, 2020.

[22] R. Katuwal, P. N. Suganthan, and L. Zhang, "Heterogeneous oblique random forest," *Pattern Recognition*, vol. 99, Article ID 107078, 2020.

[23] P. Probst, M. N. Wright, and A.-L. Boulesteix, "Hyperparameters and tuning strategies for random forest," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge*, vol. 9, no. 3, Article ID e1301, 2019.

[24] S. Kim, M. Jeong, and B. C. Ko, "Interpretation and simplification of deep forest," 2020, http://arxiv.org/abs/2001.04721.

[25] V. Jain, J. Sharma, K. Singhal, and A. Phophalia, "Exponentially weighted random forest," *Pattern Recognition And Machine Intelligence,* pp. 170–178, Lecture Notes in Computer Science, vol. 11941, Springer, Cham, Switzerland, 2019.

[26] M. StafoggiaP. Glantz et al., "A random forest approach to estimate daily particulate matter, nitrogen dioxide, and ozone at fine spatial resolution in Sweden," *Atmosphere*, vol. 11, no. 3, p. 239, 2020.

[27] S. Hauglin and P. Montesano, "Modelling above-ground biomass stock over Norway using national forest inventory data with ArcticDEM and Sentinel-2 data," *Remote Sensing of Environment*, vol. 236, p. 111501, 2020.

[28] H. R. Breidenbach and M. M. Saravi, "Land-subsidence spatial modeling using the random forest data-mining technique," *Spatial Modeling In GIS and R for Earth And Environmental Sciences*, pp. 147–159, Elsevier, Amsterdam, Netherlands, 2019.

[29] J. Cai, K. Xu, Y. Zhu, F. Hu, and L. Li, "Prediction and analysis of net ecosystem carbon exchange based on gradient boosting regression and random forest," *Applied Energy*, vol. 262, Article ID 114566, 2020.

[30] L. Wen and X. Yuan, "Forecasting CO2 emissions in Chinas commercial department, through BP neural network based on random forest and PSO," *Science of The Total Environment*, vol. 718, Article ID 137194, 2020.

[31] Y.-S. Li, H. Chi, X.-Y. Shao, M.-L. Qi, and B.-G. Xu, "A novel random forest approach for imbalance problem in crime linkage," Knowledge-Based System, Article ID 105738, 2020.

[32] S. K. Mohapatra and M. N. Mohanty, "Big data analysis and classification of biomedical signal using random forest algorithm," *New Paradigm In Decision Science And Management*, pp. 217–224, Springer, New York, NY, USA, 2020.

[33] A. Joshi, T. Choudhury, A. Sai Sabitha, and K. Srujan Raju, "Data mining in healthcare and predicting obesity," in *Proceedings of the Third International Conference on Computational Intelligence and Informatics*, pp. 877–888, Hyderabad, India, 2020.

[34] S. El-SappaghR. Sahal et al., "Alzheimer's disease progression detection model based on an early fusion of cost-effective multimodal data," *Future Generation Computer Systems*, vol. 115, pp. 680–699, 2021.

[35] Y. Saleh, A. Halidou, and P. T. Kapen, "A review of mathematical modeling, artificial intelligence and datasets used in the study, prediction and management of COVID-19," *Applied Intelligence*, vol. 50, no. 11, pp. 3913–3925, 2020.

[36] S. Khedkar, P. Gandhi, G. Shinde, and V. Subramanian, *Deep Learning and Explainable AI in Healthcare Using EHR*, pp. 129–148, Springer, New Y ork, NY, USA, 2020.

[37] T. Han, N. Stone-Weiss, J. Huang, A. Goel, and A. Kumar, "Machine learning as a tool to design glasses with controlled dissolution for healthcare applications," *Acta Biomaterials*, vol. 107, pp. 286–298, 2020.

[38] A. Subudhi, M. Dash, and S. Sabut, "Automated segmentation and classification of brain stroke using expectation-maximization and random forest classifier," *Biocybernetics and Biomedical Engineering*, vol. 40, no. 1, pp. 277–289, 2020.

[39] A. Javadi, A. Khamesipour, F. Monajemi, and M. Ghazisaeedi, "Computational modeling and analysis to predict intracellular parasite epitope characteristics using random forest technique," *Journal of Public Health*, vol. 49, no. 1, p. 125, 2020.

[40] T. Shaikhina, D. Lowe, S. Daga, D. Briggs, R. Higgins, and N. Khovanova, "Decision tree and random forest models for outcome prediction in antibody incompatible kidney transplantation," *Biomedical Signal Processing and Control*, vol. 52, pp. 456–462, 2019.

[41] K. K. Singh, S. Kumar, P. Dixit, and M. K. Bajpai, "Kalman filter based short term prediction model for COVID-19 spread," *Applied Intelligence*, pp. 1–13, 2020.

[42] S.-J. Na, J.-W. Shin, S.-H. Eom, and E.-H. Lee, "A study on random forest-based estimation model for changing the automatic walking mode of above knee prosthesis," *The Journal of IKEEE*, vol. 24, no. 1, pp. 9–18, 2020.

[43] M. Alloghani, T. Baker, D. Al-Jumeily, A. Hussain, J. Mustafina, and A. J. Aljaaf, "Prospects of machine and deep learning in analysis of vital signs for the improvement of healthcare services," *Nature-Inspired Computation In Data Mining And Machine Learning*, pp. 113–136, Springer, New York, NY, USA, 2020.

[44] Y. Zhu, W. Xu, G. Luo, H. Wang, J. Yang, and W. Lu, "Random Forest enhancement using improved Artificial Fish Swarm for the medial knee contact force prediction," *Artificial Intelligence in Medicine*, vol. 103, p. 101811, 2020.

[45] H. Zhang et al., "Deep multi-model cascade method based on CNN and random forest for pharmaceutical particle detection," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 9, pp. 7028–7042, 2020.

[46] H. Lee and E. Jung, *"An Analysis of Annual Changes on the Determining Factors for Teacher Attachment with Random Forest*, pp. 463–470, Springer, New York, NY, USA, 2020.

[47] X. LiuL. Liu et al., "Downscaling of solar-induced chlorophyll fluorescence from canopy level to photosystem level using a random forest model," *Remote Sensing of Environment*, vol. 231, Article ID 110772, 2019.

[48] S. Guanter and J. Santosh Kumar, "Performance evaluation of random forest with feature selection methods in prediction of diabetes," *International Journal of Electrical and Computer Engineering*, vol. 10, 2020.

[49] A. Subasi, A. Ahmed, E. Aličković, and A. Rashik Hassan, "Effect of photic stimulation for migraine detection using random forest and discrete wavelet transform," *Biomedical Signal Processing and Control*, vol. 49, pp. 231–239, 2019.

[50] N. El Haouij, J.-M. Poggi, R. Ghozi, S. Sevestre-Ghalila, and M. Jaïdane, "Random forest-based approach for physiological functional variable selection for driver's stress level classification," *Statistical Methods & Applications*, vol. 28, no. 1, pp. 157–185, 2019.

[51] D. Ayata, Y. Yaslan, and M. E. Kamasak, "Emotion recognition from multimodal physiological signals for emotion aware healthcare systems," *Journal of Medical and Biological Engineering*, vol. 40, pp. 149–157, 2020.

[52] M. Zeraatpisheh, E. Bakhshandeh, M. Hosseini, and S. M. Alavi, "Assessing the effects of deforestation and intensive agriculture on the soil quality through digital soil mapping," *Geoderma*, vol. 363, Article ID 114139, 2020.

[53] X. Du, P. Wang, L. Fu, H. Liu, Z. Zhang, and C. Yao, "Determination of chlorpyrifos in pears by Raman spectroscopy with random forest regression analysis," *Analytical Letters*, vol. 53, no. 6, pp. 821–833, 2020.

[54] J. Wang, R. Zuo, and Y. Xiong, "Mapping mineral prospectivity via semi-supervised random forest," *Natural Resources Research*, vol. 29, no. 1, pp. 189–202, 2020.

[55] S. Liu, X. Qian, H. Wan, Z. Ye, S. Wu, and X. Ren, "NPC three-level inverter open-circuit fault diagnosis based on adaptive electrical period partition and random forest," *Journal of Sensor and Actuator Networks*, vol. 2020, Article ID 9206579, 18 pages, 2020.

[56] S. S. Rathore and S. Kumar, "An empirical study of ensemble techniques for software fault prediction," *Applied Intelligence*, pp. 1–30, 2020.

[57] T. Ahmad and H. Chen, "Nonlinear autoregressive and random forest approaches to forecasting electricity load for utility energy management systems," *Sustainable Cities and Society*, vol. 45, pp. 460–473, 2019.

[58] S. Gupta, J. Sarkar, M. Kundu, N. R. Bandyopadhyay, and S. Ganguly, "Automatic recognition of SEM microstructure and phases of steel using LBP and random decision forest operator," *Measurement*, vol. 151, Article ID 107224, 2020.

[59] L. T. T. Ho, L. Dubus, M. De Felice, and A. Troccoli, "Reconstruction of multidecadal country-aggregated hydro power generation in Europe based on a random forest model," *Energies*, vol. 13, no. 7, p. 1786, 2020.

[60] Y. Zhou, S. Li, C. Zhou, and H. Luo, "Intelligent approach based on random forest for safety risk prediction of deep foundation pit in subway stations," *Journal of Computing in Civil Engineering*, vol. 33, no. 1, Article ID 05018004, 2019.

[61] X. DengY. Zhan et al., "Predictive geographical authentication of green tea with protected designation of origin using a random forest model," *Food Control*, vol. 107, Article ID 106807, 2020.

[62] S. A. Liu, P. Ngare, and D. Ikpe, "Probabilistic forecasting of crop yields via quantile random forest and Epanechnikov

Kernel function," *Agricultural and Forest Meteorology*, vol. 280, Article ID 107808, 2020.

[63] H. J. Jeong and M. H. Kim, "Utilizing adjacency of colleagues and type correlations for enhanced link prediction," *Data & Knowledge Engineering*, vol. 125, Article ID 101785, 2020.

[64] Z. Khorshidpour, S. Hashemi, and A. Hamzeh, "Evaluation of random forest classifier in security domain," *Applied Intelligence*, vol. 47, no. 2, pp. 558–569, 2017.

[65] J. Tian, L. Liu, F. Zhang, Y. Ai, R. Wang, and C. Fei, "Multi-domain entropy-random forest method for the fusion diagnosis of inter-shaft bearing faults with acoustic emission signals," *Entropy*, vol. 22, no. 1, p. 57, 2020.

[66] B. Shaw, A. K. Suman, and B. Chakraborty, *Wine Quality Analysis Using Machine Learning*, pp. 239–247, Springer, New York, NY, USA, 2020.

[67] K. Madhumathi and T. Suresh, *Node Localization in Wireless Sensor Networks Using Multi-Output Random Forest Regression*, pp. 177–186, Springer, New York, NY, USA, 2020.

[68] Y. Fang, Y. Xu, C. Huang, L. Liu, and L. Zhang, "Against malicious SSL/TLS encryption: identify malicious traffic based on random forest," in *Proceedings of the Fourth International Congress on Information And Communication Technology*, pp. 99–115, London, UK, 2020.

[69] T. T. Bhavani, M. K. Rao, and A. M. Reddy, "Network intrusion detection system using random forest and decision tree machine learning techniques," in *Proceedings of the First International Conference On Sustainable Technologies For Computational Intelligence*, pp. 637–643, London, UK, 2020.

[70] P. S. Chaithanya, M. R. G. Raman, S. Nivethitha, K. S. Seshan, and V. S. Sriram, "An efficient intrusion detection approach using enhanced random forest and moth-flame optimization technique," *Computational Intelligence In Pattern Recognition*, pp. 877–884, Springer, New York, NY, USA, 2020.

[71] Z. Mingjing, "A novel strategy for quantitative analysis of soil pH via laser-induced breakdown spectroscopy coupled with random forest," *Plasma Science Technology*, vol. 22, no. 7, p. 74003, 2020.

[72] M.-H. Lee, "Robust random forest based non-fullerene organic solar cells efficiency prediction," *Organic Electronics*, vol. 76, Article ID 105465, 2020.

[73] J. Zhang, G. Ma, Y. Huang, J. sun, and F. Aslani, "Modelling uniaxial compressive strength of lightweight self-compacting concrete using random forest regression," *Construction and Building Materials*, vol. 210, pp. 713–719, 2019.

[74] W. Nener, C. Wu, H. Zhong, Y. Li, and L. Wang, "Prediction of undrained shear strength using extreme gradient boosting and random forest based on Bayesian optimization," *Geoscience Frontiers*, vol. 12, no. 1, pp. 469–477, 2020.

[75] P. Zhang, Z.-Y. Yin, Y.-F. Jin, and T. H. T. Chan, "A novel hybrid surrogate intelligent model for creep index prediction based on particle swarm optimization and random forest," *Engineering Geology*, vol. 265, p. 105328, 2020.

[76] K. Shah, H. Patel, D. Sanghvi, and M. Shah, "A comparative analysis of logistic regression, random Forest and KNN models for the text classification," *Augmented Human Research*, vol. 5, no. 1, pp. 1–16, 2020.

[77] S. Hakak, M. Alazab, S. Khan, T. R. Gadekallu, P. K. R. Maddikunta, and W. Z. Khan, "An ensemble machine learning approach through effective feature extraction to classify fake news," *Future Generation Computer Systems*, vol. 117, pp. 47–58, 2021.

[78] S. N. Singh and T. Sarraf, "Sentiment analysis of a product based on user reviews using random forests algorithm," *Data Science & Engineering*, vol. 32, pp. 112–116, 2020.

[79] H. Zhang, N. Cheng, Y. Zhang, and Z. Li, "Label flipping attacks against Naive Bayes on spam filtering systems," *Applied Intelligence*, 2021.

[80] R. P. Kaur, M. Kumar, and M. K. Jindal, "Newspaper text recognition of Gurumukhi script using random forest classifier," *Multimedia Tools and Applications Journal*, pp. 1–14, 2019.

[81] S. Madichetty and M. Sridevi, "A novel method for identifying the damage assessment tweets during disaster," *Futur. Gener. Comput. Syst.*vol. 116, pp. 440–454, 2020.

[82] A. Madasu and S. Elango, "Efficient feature selection techniques for sentiment analysis," *Multimedia Tools and Applications*, vol. 79, no. 9-10, pp. 6313–6335, 2020.

[83] A.-C. Chang, C. V. Trappey, A. J. C. Trappey, and L. W. L. Chen, "Web mining customer perceptions to define product positions and design preferences," *International Journal on Semantic Web and Information Systems*, vol. 16, no. 2, pp. 42–58, 2020.

[84] R. Kumar and J. Kaur, "Random forest-based sarcastic tweet classification using multiple feature collection," in *Multimedia Big Data Computing For IoT Applications*, pp. 131–160, Springer, New York, NY, USA, 2020.

[85] A. Onan and M. A. Toçouglu, "Weighted word embeddings and clustering-based identification of question topics in MOOC discussion forum posts," *Computer Applications in Engineering Education*, 2020.

[86] O. M. Baez-Villanueva and M. Zambrano, "RF-MEP: a novel Random Forest method for merging gridded precipitation products and ground-based measurements," *Remote Sensing of Environment*, vol. 239, Article ID 111606, 2020.

[87] A. Beck, "Sentiment analysis on massive open online course evaluations: a text mining and deep learning approach," *Computer Applications in Engineering Education*, 2020.

[88] A. Onan, "Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks," *Computer Applications in Engineering Education*, Article ID e5909, 2020.

[89] A. Onan and M. A. Tocoglu, "A term weighted neural language model and stacked bidirectional LSTM based framework for sarcasm identification," *IEEE Access*, vol. 9, pp. 7701–7722, 2021.

[90] C. Rodriguez-Pardo, M. A. Patricio, A. Berlanga, and J. M. Molina, *"Machine Learning for Smart Tourism and Retail*, pp. 311–333, IGI Global, 2020.

[91] W. Song and Y. Zhou, "Road travel time prediction method based on random forest model," in *Smart Trends In Computing And Communications*, pp. 155–163, Springer, New York, NY, USA, 2020.

[92] A. Jamatia, U. Baidya, S. Paul, S. DebBarma, and S. Dey, "Rating prediction of tourist destinations based on supervised machine learning algorithms," *Computational Intelligence In Data Mining*, pp. 115–125, Springer, New York, NY, USA, 2020.

[93] C. F. Baumeister, T. Gerstenberg, T. Plieninger, and U. Schraml, "Exploring cultural ecosystem service hotspots: linking multiple urban forest features with public participation mapping data," *Urban Forestry & Urban Greening*, vol. 48, Article ID 126561, 2020.

[94] J. Evans, B. Waterson, and A. Hamilton, "Forecasting road traffic conditions using a context-based random forest

algorithm," *Transportation Planning and Technology*, vol. 42, no. 6, pp. 554–572, 2019.

[95] L. Zhou, X. Dang, Q. Sun, and S. Wang, "Multi-scenario simulation of urban land change in Shanghai by random forest and CA-Markov model," *Sustainable Cities and Society*, vol. 55, Article ID 102045, 2020.

[96] H. Liang, Z. Guo, J. Wu, and Z. Chen, "GDP spatialization in Ningbo City based on NPP/VIIRS night-time light and auxiliary data using random forest regression," *Advances in Space Research*, vol. 65, no. 1, pp. 481–493, 2020.

[97] Z. Mei, W. Ding, C. Feng, and L. Shen, "Identifying commuters based on random forest of smartcard data," *IET Intelligent Transport Systems*, vol. 14, no. 4, pp. 207–212, 2020.

[98] Q. Li, L. Chen, X. Li et al., "A progressive random forest-based random walk approach for interactive semi-automated pulmonary lobes segmentation," *International Journal of Machine Learning and Cybernetics*, vol. 11, pp. 2221–2235, 2020.

[99] S. L. S. Darshan and C. D. Jaidhar, "An empirical study to estimate the stability of random forest classifier on the hybrid features recommended by filter based feature selection technique," *International Journal of Machine Learning and Cybernetics*, vol. 11, no. 2, pp. 339–358, 2020.

[100] S. K. Lakshmanaprabu, K. Shankar, M. Ilayaraja, A. W. Nasir, V. Vijayakumar, and N. Chilamkurti, "Random forest for big data classification in the internet of things using optimal features," *International Journal of Machine Learning and Cybernetics*, vol. 10, no. 10, pp. 2609–2618, 2019.

[101] P. Liu, X. Wang, L. Yin, and B. Liu, "Flat random forest: a new ensemble learning method towards better training efficiency and adaptive model size to deep forest," *International Journal of Machine Learning and Cybernetics*, vol. 11, pp. 2501–2513, 2020.

[102] M. Goodwin and A. Yazidi, "Distributed learning automata-based scheme for classification using novel pursuit scheme," *Applied Intelligence*, vol. 50, no. 7, pp. 2222–2238, 2020.

[103] K. S. Narendra and M. A. L. Thathachar, *Learning Automata: An Introduction*, Courier Corporation, Chelmsford, CA, USA, 2012.

[104] A. Rezvanian, A. M. Saghiri, S. M. Vahidipour, M. Esnaashari, and M. R. Meybodi, *Recent Advances In Learning Automata*, Vol. vol. 754, Springer, New York, NY, USA, 2018.

[105] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2017.

[106] B. Pang and L. Lee, "Seeing stars: exploiting class relationships for sentiment categorization with respect to rating scales," 2005, http://arxiv.org/abs/0506075.

[107] A. Maas, R. E. Daly, P. T. Pham, D. Huang, A. Y. Ng, and C. Potts, "Learning word vectors for sentiment analysis," in *Proceedings Of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pp. 142–150, Buffalo, NY, USA, 2011.

[108] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? sentiment classification using machine learning techniques," 2002, http://arxiv.org/abs/0205070.

[109] D. Dua and C. Graff, *UCI Machine Learning Repository*, University of California, School of Information and Computer Science, Irvine, CA, 2019.

[110] C. López-Vázquez and E. Hochsztain, "Extended and updated tables for the Friedman rank test," *Communications in Statistics - Theory and Methods*, vol. 48, no. 2, pp. 268–281, 2019.