Soil Classification From Large Imagery Databases Using a Neuro-Fuzzy Classifier
Classification de sols à partir de larges bases de données d'imagerie utilisant un classificateur Neuro-Flu

Soumadip Ghosh, Debasish Biswas, Sushanta Biswas, Debasree Chanda (Sarkar), and Partha Pratim Sarkar

***Abstract*— In this paper, we propose a neuro-fuzzy (NF) classification technique to determine various soil classes from large imagery soil databases. The technique looks at the feature-wise degree of belongings of the imagery databases to obtainable soil classes using a fuzzification method. The fuzzification method builds a membership matrix with an element count equal to the mathematical product of the number of data records and soil classes present. The elements of this matrix are the input to a neural network model. We apply our technique to three UCI databases, namely, Statlog Landsat Satellite, Forest Covertype, and Wilt for soil classification. The paper aims to find out soil classes using the proposed technique, and then compare its performance with four well-known classification algorithms, namely, radial basis function network, k-nearest neighbor, support vector machine, and adaptive NF inference system. Numerous measures, for example, root-mean-square error, kappa statistic, accuracy, false positive rate, true positive rate, precision, recall, F-measure, and area under the curve, are used for evaluating the quantitative analysis of the simulated results. All these evaluation measures approve the supremacy of the proposed NF method.**

***Résumé*— Dans cet article, nous proposons une technique de classification neuro-floue (NF) pour déterminer des classes de sol variées à partir de larges bases de données d'imagerie de sol. La technique regarde la caractéristique-sage de degré d'appartenance des bases de données d'imagerie à des classes de sol pouvant être obtenus en utilisant une méthode de fuzzification. La méthode de fuzzification construit une matrice d'appartenance avec un élément compte égal au produit mathématique du nombre d'enregistrements de données et les classes de sol présentes. Les éléments de cette matrice sont l'entrée d'un modèle de réseau de neurones. Nous appliquons notre technique à trois bases de données de l'UCI, à savoir, Statlog Landsat Satellite, Forest Covertype et Wilt pour la classification des sols. L'article vise à trouver des classes de sol en utilisant la technique proposée, puis comparer ses performances avec quatre algorithmes de classification bien connus, à savoir, les réseaux de neurones à fonction de base radiale, le k-plus proche voisin, la machine à vecteurs de support, et le système d'inférence NF adaptative. De nombreuses mesures, par exemple, l'erreur quadratique moyenne, le kappa statistique, la précision, le taux de faux positifs, le taux de vrais positifs, la précision, le rappel, F-mesure, et l'aire sous la courbe, sont utilisées pour l'évaluation de l'analyse quantitative des résultats de simulation. Toutes ces mesures d'évaluation approuvent la supériorité de la méthode NF proposée.**

***Index Terms*— Fuzzy neural networks, geographic information systems, image classification, image databases, information retrieval.**

I. Introduction

DATA mining techniques typically analyze large imagery databases and set up useful classification and patterns to develop geographical information systems (GIS)-based frameworks for industrial, scientific, and commercial purposes [1], [2]. Research studies in the data mining field have used various techniques of data analysis, including decision trees, genetic algorithms, machine learning, and other statistical analysis methods [3], [4] for building such GIS-based frameworks. Large imagery soil databases, covering a vast geographical area, provide a distinctive prospect for developing land use and land cover information describing several soil classes. Regularly updated land use and land cover information is very much crucial to manyenvironmental and socioeconomic applications based on GIS, comprising regional and urban planning, conservation and management of natural resources, and so on. The research objective leads to a new area of data mining referred to as soil data mining. The present work outlines a research study involving useful data mining techniques to improve the effectiveness and classification accuracy of massive imagery soil databases.

Manuscript received November 18, 2015; revised February 15, 2016 and May 20, 2016; accepted July 20, 2016. Date of current version December 13, 2016.

S. Ghosh is with the Academy of Technology, Kolkata 712121, India (e-mail: soumadip.ghosh@gmail.com).

D. Biswas is with the Dr. Meghnad Saha Institute of Technology, Haldia 721657, India (e-mail: debasish79@gmail.com).

S. Biswas, D. Sarkar, and P. P. Sarkar are with the Department of Engineering and Technological Studies, University of Kalyani, Kalyani 741235, India (e-mail: biswas.su@gmail.com; dsarkar70@gmail.com; parthabe91@yahoo.co.in).

Associate Editor managing this paper's review: April Khademi.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/CJECE.2016.2596767

General information related to land comes from soil survey analysis. It is the procedure of categorizing various soil types or using other attributes of the ground cover over a particular geographical area for soil mapping [5]. Field sampling usually constructs the major data for soil survey analysis. The goal is to perform categorization of large imageries in soil mapping. This kind of information can be very much beneficial for the researchers and knowledge workers for performing the analysis. Classification [6] is an important data mining technique typically used to perform soil survey for building GIS-based frameworks. Thus, soil classification as a form of soil survey analysis can analyze large imagery databases and develop worthwhile knowledge for GIS-based frameworks. That is why such an analysis is essential for the research field related to soil data mining. For producing updated land use and land cover information at diverse scales, various classification techniques have been established.

Artificial neural network (ANN) [7]–[9] is a prevailing modeling tool that can achieve human-like rational thinking. It is widely known for excellent precision and extraordinary learning capability even when negligible facts are accessible. One of the proficient methods of classification from the ANN domain is the multilayer perceptron (MLP) model [10], [11]. An MLP model contains several layers of nodes arranged in a directed graph structure, with connections between adjacent layers. MLP uses the backpropagation technique to train the network. The radial basis function network (RBFN) [12], [13] is another influential model that uses radial basis function (RBF) as the activation function. The output of RBFN is the linear combination of RBFs of inputs and neuron parameters. RBFNs are typically used in system control, approximation of functions, classification, and prediction of time series.

The k-nearest neighbor (k-NN) [14] is an instance-based learning method for classifying objects using the rationale of nearest training examples within the search space. The procedure compares a given test pattern with training patterns that are similar to it. One should use suitable distance metric to assign a new data point to the most frequently occurring class in the neighborhood. The distance metric used in this approach are the Euclidean distance for continuous valued variables or the Hamming distance for discrete valued variables. The method works well in classifying the imagery databases where the data reveal spatial properties.

Support vector machine (SVM) [15] is another powerful supervised learning model used in the machine learning field. It constructs one or more than one hyperplane in a high-dimensional feature space for regression, classification, or some other analysis tasks. SVM employs a hyperplane to differentiate between classes. When classes overlap, one should use hyperplanes to minimize the error of data points alongside the boundary line among classes. These points are the support points or support vectors.

However, no matter how good a classifier is, several uncertainties or ambiguities can still rise at any period of a classification procedure. It occurs due to the existence of vagueness in input data, intersecting boundaries between several classes, and incomprehensibility in describing features. The fuzzy set theory [16]–[18] is flexible enough to deal with the several aspects of indecisiveness about real-life circumstances. ANN combined with the fuzzy set theory-based method is the neuro-fuzzy (NF) technique [19]–[21]. The hybrid approach combines the human-like rational thinking of fuzzy-based models with the wisdom and connection-oriented structure of ANNs using the fuzzy sets and linguistic model-based procedures.

There is one variant of NF model named adaptive NF inference system (ANFIS) [19]. ANFIS is a powerful classification model that generates a set of interpretable IF–THEN rules. The nodes of an adaptive neural network are associated with certain parameters that will decide the final output. It normally employs a hybrid learning algorithm combining gradient descent and least square methods to adjust the neural network parameters efficiently in an adaptive network.

In our work, we propose an NF system (NFS) for soil data mining. The proposed method considers the feature-wise degree of memberships of large imageries to existing soil classes that are performed using a fuzzification process. The process produces a membership matrix with an element count equal to the product of number of data records and classes present. These matrix elements are the input to an ANN model. We apply our method to three UCI databases, namely, Statlog Landsat Satellite, Forest Covertype, and Wilt data sets for soil classification and, therefore, compare its performance with RBFN, k-NN, SVM, and ANFIS models. We investigate the performances of these classifiers using different measures such as root-mean-square error (RMSE), accuracy, false positive (FP) rate, true positive (TP) rate, kappa statistic, recall, precision, F-measure, and area under the curve (AUC).

The paper is organized as follows. Section II consists of the related works done in this field. Section III describes the proposed NF classification method. Section IV depicts the detailed procedure in terms of the proposed NF method, RBFN, k-NN, SVM, and ANFIS. Section V explains the performance analysis and results, and Section VI is attributed to the conclusion.

II. Related Works

In the present context, we discuss some of the notable research works done in the field of soil classification from large imagery databases.

Shine and Carr [22] offered significant contributions to this field. They envisioned the statistical relations between land covers and spatial compression procedures for mapping and categorizing high-resolution digital satellite imageries. They conducted their research using a multispectral imagery of New York. As the data set was massive in size, they used lossless compression techniques for reducing computational complexity. The research study significantly contributed to the field of soil classification domain. Besides this, the research work was an elongation of [23] and [24].

Later, Shine and Carr [25] developed useful techniques for classification of satellite imageries for soil mapping. They employed different classification methods to the Statlog Landsat imaging database of Australia and the multispectral database of the south-central Virginia. These classifiers were classification tree (CT), ANN, discriminant analysis (DA), k-NN, and SVM. According to the performance, the k-NN performed the highest; then in precision were CT, ANN, SVM, and DA successively.

Rozenstein and Karnieli [26] created a low-cost GIS model. They combined the native GIS data with the satellite remote sensing data to generate a moderately accurate land map of northern Negev region in Israel. Besides this, additional land-use data were used to improve the remote sensing classification accuracy inside the GIS framework. The research work also integrated unsupervised and supervised learning methods, which resulted in a classification accuracy of 81%.

Lu and Weng [27] considered classifying large-scale imagery data sets using diverse aspects of the image classification methods used. Their work showed that using efficient utilization of multiple characteristics of a large imaging data set, the accuracy improved significantly.

Alajlan *et al*. [28] designed a framework to enhance the classification performance of hyperspectral images by combining supervised and unsupervised learning approaches. Basically, they combined SVM classifier with fuzzy C-means clustering algorithm in their work. SVM classifier was used to generate a spectral-based classification map, and the fuzzy C-means clustering was applied to provide an ensemble of clustering maps. The experimental results confirmed the promising capabilities of their proposed framework.

Li *et al*. [29] studied various remotely sensed image classifiers, involving object-based, pixel-wise, and subpixel-wise image classifiers and emphasized the significance of including spatial and contextual based information in remote sensing image classification. Furthermore, this study grouped the spatial and contextual analysis techniques into three broad groups, containing: 1) texture mining; 2) Markov random fields modeling; and 3) image segmentation and object-based image analysis. The work also explained the need to develop and analyze geographical information models for spatial and contextual classifications using some specific case studies.

III. Proposed Neuro-Fuzzy Classification Method

In this paper, we propose an NF classification method to identify different soil types from large imagery databases. The idea is to consider the feature-wise degree of memberships of such imagery databases to the soil classes that are done using a fuzzification process. The fuzzification process is based on the concept of generating a membership matrix as specified in [30]. The work designed an NF procedure that could build a matrix with a total number of elements equal to the multiplication product of the number of data records and obtainable classes. This NF method denoted a generalized classification technique and was applied to various application areas. Basically, it employed the generalized bell-shaped membership function (MF) with three parameters for the fuzzification purpose.

In the area of soil classification using imageries, researchers are still in search of a procedure that can provide an enhanced performance. Essentially, the proposed NF technique here is tuned to provide optimized performance for soil classification from large imagery databases. The MF used in this paper is completely different in nature and functionality from the previous one. The proposed hybrid classifier uses the *Gaussian* MF with two parameters for fuzzification. This Gaussian MF is mainly specialized for use in the image classification domain. As usual, the fuzzification process gives membership values for each data pattern in the database to all existing classes to form a membership matrix. The elements of this matrix are the input to an MLP model (i.e., ANN). The number of output neurons of the ANN equals the number of classes present. Defuzzification process then applies to the ANN model output. A hard classification [19] of the input data patterns is achieved using a maximum operation on the output of ANN as in the case of a traditional fuzzy classification-based system. The proposed method is divided into three different phases described in the following.

A. First Phase (Fuzzification)

The first phase named *fuzzification* takes a sample data set containing various data patterns, fuzzifies the data pattern values with a Gaussian MF, and then computes the degree of membership of individual data patterns to numerous classes. Let us consider that we have a data set consisting of B input patterns and A data classes. We now define the data set in terms of the input pattern vector $\boldsymbol{x}$ as

$$x = [x_1, x_2, \ldots, x_B]^T \quad (1)$$

provided that T is the matrix transpose operator.

The fuzzification phase mainly constructs a membership matrix of order $(B \times A)$ from the input pattern vector $\boldsymbol{x}$ that consists of the degree of the memberships of B different patterns to A number of classes. Each and every element in the matrix is an MF of the form $\boldsymbol{g_{u,v}}$ $(\boldsymbol{x_u})$, where x_u is the uth feature value of the input pattern vector $\boldsymbol{x}$ with indices $u = 1, 2, \ldots, B$ and $v = 1, 2, \ldots, A$. Thus, we can describe the MF as follows:

$$g_{u,v}(x_u) = \text{degree of membership of pattern } u \text{ with respect to class } v. \quad (2)$$

As already stated, we have used the well-known Gaussian MF for fuzzification. The Gaussian curve MF has a smooth curvature and nonzero at all points. It is also symmetric in nature and depends upon two different parameters σ and c as given by the following equation:

$$g(x_u; \sigma, c) = e^{\frac{-(x_u - c)^2}{2\sigma^2}}. \quad (3)$$

Basically, the parameters c and σ here represent the Gaussian curve MF's center and width, respectively. They can control the shape and curvature of the MF. If we modify the values of these parameters, we can get the desired MF that offers more flexibility for classification. The MF is easier to model and train in an ANN model. Besides this, it is suitable for performing image classification using imagery databases. Thus, an NF model with a Gaussian MF is proposed and used in this paper that is more straightforward to implement in the neural network model. The Gaussian curve MF used here is shown in Fig. 1.

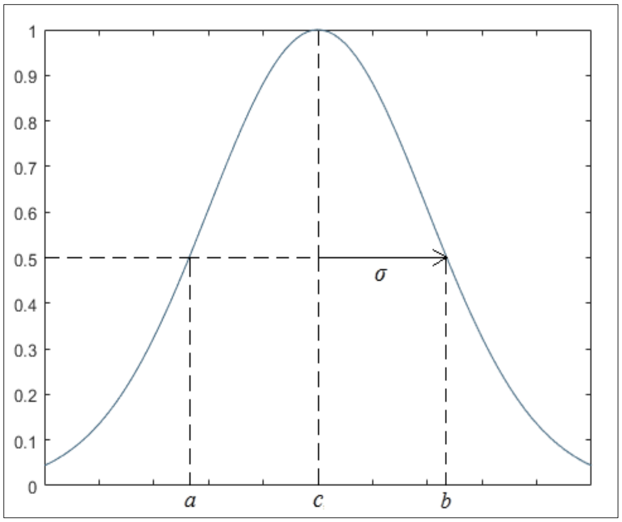


Fig. 1. Gaussian curve MF used in fuzzification.

The Gaussian curve MF presented in Fig. 1 has a center at c, with $c = (a + b)/2.0$, where a and b are the two crossover points in the curve. Here, the curve is symmetric in nature, and therefore $\sigma = |b - c| = |c - a|$. The membership value at the crossover points is 0.5, and at the center c, its value is 1.0 (i.e., maximum). The selection of membership value is done in a way so that every training pattern acquires a membership value of 1.0 when it is at the center of the MF, and when it is moved away from the center, its value gradually drops and reaches 0.5 at the boundary of the training data set.

The center position c is evaluated as the mean (i.e., average) value of the training data set. It is selected as c = mean_val(x) (i.e., the mean value of the data set for a pattern x). The two crossover points a and b in the MF curvature are calculated as a = mean_val(x) − [max_val(x)-min_val(x)]/*2.0* and b = mean_val(x) + [max_val(x) -min_val(x)]/*2.0*, where min_val and max_val are the minimum and maximum values, respectively, of the data set for a specific data pattern x. Such a selection of a and b confirms that most of the training data patterns will have membership values ≥0.5, and test data patterns will have membership values in the interval [0, 1].

We apply the above MF to the input pattern vector x for constructing the membership matrix. The resultant matrix has the following structure:

$$G(x) = \begin{bmatrix} g_{1,1}(x_1) & g_{1,2}(x_1) & g_{1,3}(x_1) & \cdots & g_{1,A}(x_1) \\ g_{2,1}(x_2) & g_{2,2}(x_2) & g_{2,3}(x_2) & \cdots & g_{2,A}(x_2) \\ g_{3,1}(z_3) & g_{3,2}(x_1) & g_{3,3}(x_3) & \cdots & g_{3,A}(x_3) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ g_{B,1}(x_B) & g_{B,2}(x_B) & g_{B,3}(x_B) & \cdots & g_{B,A}(x_B) \end{bmatrix}. \tag{4}$$

Here, $\boldsymbol{g_{u,v}(x_u)}$ is the degree of belongings of the uth data pattern in the input vector $\boldsymbol{x}$ to the class v using $u = 1, 2, \ldots, B$ and $v = 1, 2, \ldots, A$. For example, $\boldsymbol{m}_{2,4}(\boldsymbol{x}_2)$ denotes the degree of membership of the second pattern to class 4. The membership matrix is used as the input to an ANN model as presented in the following.

TABLE I

CONFIGURATION PARAMETERS OF THE MLP MODEL

Parameter	**Value**
Number of hidden layers	two
Number of neurons in input layer	element count in membership matrix
Number of neurons in output layer	data classes present
Learning rule	gradient descent with momentum
Transfer function used	tan-sigmoid

B. Second Phase (Development of the ANN Model)

The second phase constructs an MLP model (ANN). In this phase, the membership matrix mentioned above is transformed into an $(A \times B)$ vector by performing transpose operation on all tuples and attributes. The newly generated vector is the input to an MLP classifier. It is seen that using more than one hidden layer in the ANN model gives better result in image classification domain. Thus, this paper uses an MLP model that consists of one input layer, one output layer, and *two hidden layers* in between them. As usual, the number of neurons present in the two hidden layers of the current ANN model is also different. However, [30] used a single hidden layer in its ANN model. Table I provides the list of configuration parameters used in the given MLP model.

The selection of the number of neurons obtainable in the hidden layers is too an important constraint. Subsequently, a thorough investigation helps us in selecting the number of PEs present in the hidden layers [31]. The first hidden layer contains some PEs, denoted by h_1, and is given by the following equation:

$$h_1 = \sqrt{(\text{number of inputs} * \text{number of outputs})}. \tag{5}$$

The second hidden layer contains a few number of neurons, represented by h_2, as specified by

$$h_2 = \frac{2}{3} * (\text{number of inputs} + \text{number of outputs}). \tag{6}$$

C. Third Phase (Defuzzification)

The third or final phase employs the *defuzzification* procedure that is theoretically just opposite to the first phase. In this phase, the proposed classification model implements a hard classification using a maximum operation to create the activation output of the MLP model. An input pattern will be associated with a particular *class t* provided that the pattern has the maximum class membership value with respect to class t compared with other classes. Therefore, an unknown pattern $\boldsymbol{z}$ is assigned to the *class t* based on the concept of "maximum class membership value" if and only if

$$G_t(z) \geq G_w(z) \quad \forall w \in (1, 2, \ldots, A) \text{ and } w \neq t \tag{7}$$

where $\boldsymbol{G_w(z)}$ is the activation output of the wth node in the last layer (i.e., output layer) of the RBFN model.

The block diagram of the proposed system model is shown in Fig. 2.

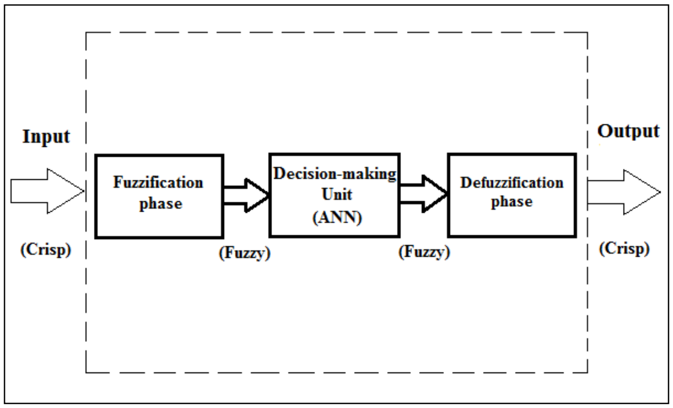


Fig. 2. Proposed NF classification system model.

IV. Detailed Procedure

Basically, classification procedure is divided into two major steps; in the first step, a classification model is developed indicating a well-defined set of classes. Therefore, this is the *training step*, where a classification technique constructs the model from a training data set. After that, the model is used for prediction called the *testing step* where the accuracy of the model is estimated using a test data set independent of the training data set. Some preprocessing techniques like data cleaning and data transformation are to be applied before the actual classification procedure. This paper makes use of the NFS, RBFN, k-NN, SVM, and ANFIS classifiers using three large imagery soil data sets. The broad-level stages of the procedure are described in the following.

Stage 1) These preprocessing techniques are applied to each data set before the classification task.

Stage 1a) *Data Cleaning:* The process simply substitutes the missing values for an attribute by its arithmetic mean.

Stage 1b) *Data Transformation:* The method normalizes a data set by transforming attribute values to a specified interval $[-1.0, +1.0]$. It is applicable to ANN-based methods only.

Stage 2) Afterward, every single data set is distributed into two separate subsets, namely, the training set and the test set. The study employs *tenfold cross-validation* [32] technique to generate the training and test data sets so that they should be independent of each other for avoiding biases. Therefore, the generated training and test data sets are entirely disjoint. However, [30] used predetermined two-third of the data set for training and the remaining one-third data for testing, which might allow biasness to creep in. Therefore, the proposed NF method here can deliver a more accurate classification performance compared with the previous technique because it does not suffer from the problem of overfitting.

Stage 3) The proposed NFS technique employs the training set for building a classification model. The training set is also specified to RBFN, k-NN, SVM, and ANFIS techniques independently for building other models.

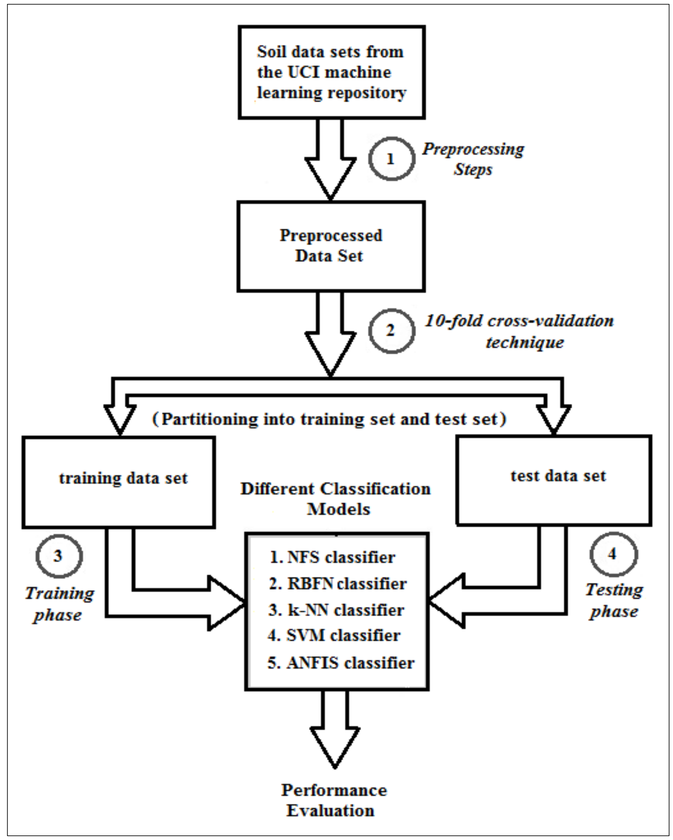


Fig. 3. Broad-level stages of the detailed procedure.

Stage 4) The five classification models (NFS, RBFN, k-NN, SVM, and ANFIS) later applied to the test data set for performance evaluation using measures such as RMSE, accuracy, FP rate, TP rate, kappa statistic, recall, precision, F-measure, and AUC.

The broad-level stages of the complete procedure are depicted in Fig. 3.

V. Results and Discussion

The five classification models, namely, NFS, RBFN, k-NN, SVM, and ANFIS, are trained and tested on three UCI soil data sets using the MATLAB software (version R2015a). After building the models, they are employed to the test data set for performance evaluation. So far, this paper has discussed about the configuration of NFS model only. Section V-A here describes the configurations of the other classification models used in the simulation. Section V-B denotes the various evaluation measures, and Section V-C provides the simulation results and performance analysis.

A. Configurations of Other Models

1) *SVM:* The study uses an SVM model with a Gaussian RBF kernel for image classification. A nonlinear version of SVM can be represented using a kernel function K as

$$K(x_i \cdot x_j) = \phi(x_i) \cdot \phi(x_j). \tag{8}$$

Here $\varphi(x)$ is the nonlinear mapping function employed to map the data tuples in the imagery database.

An SVM model with a Gaussian RBF kernel is defined as

$$K(x_i \cdot x_j) = e^{\frac{-\|x_i - x_j\|^2}{2\sigma^2}}. \tag{9}$$

2) *k-NN:* The instance-based learning model here uses the Mahalanobis distance as the distance metric to select the nearest neighbors.
3) *RBFN:* The RBFN model employs a Gaussian RBF activation function in the hidden layer unit. It uses the fuzzy C-means algorithm to find out the RBF centers.
4) *ANFIS:* It is a Sugeno-type fuzzy inference system model. The classifier applies a hybrid learning algorithm combining backpropagation gradient descent and least-square methods to regulate the neural network parameters.

B. Performance Measures

We estimate the performances of these models using various evaluation metrics, such as RMSE [33], kappa statistic [34], and various measures resulting from the confusion matrix [35]. The confusion matrix measures are typically accuracy, FP rate, TP rate, precision, recall, F-measure, and AUC [6] values.

C. Results and Performance Analysis

We apply NFS, RBFN, k-NN, SVM, and ANFIS classifiers to three UCI imagery data sets for performance analysis as described in the following. The results reported here are solely based on the simulation experiment that we have taken.

1) Statlog Landsat Satellite Database: We use the UCI Statlog Landsat Satellite [36] database of agricultural land in Australia to classify the different soil classes constituting dissimilar soil types. The database was built considering only a small division (82 rows $\times$ 100 columns) of the original Landsat multispectral scanner imaging data set. It has four spectral bands in a single image frame. In a satellite image, 3×3 (=9) square neighborhood of pixels was designated, and the concerning four spectral values of pixels were computed. This multivariate data set consists of 6435 tuples, 36 (=9 pixels in the neighborhood $\times$ 4 spectral bands) input features, and one class attribute. Each of the input attributes is quantitative in nature, and the value should lie inside 0-255. This paper has used four features (attributes 17, 18, 19, and 20) only as suggested by the UCI (i.e., the four spectral values for the central pixel).

The class label attribute comprises six values indicating six categories of soil. They are, namely, red soil (class 1), cotton crop soil (class 2), gray soil (class 3), damp gray soil (class 4), vegetation stubble soil (class 5), and very damp gray soil (class 7). There is no example of category 6 (mixture type) in the database. The image of the Statlog Landsat Satellite database is shown in Fig. 4.

Each of the five classifiers, namely, NFS, RBFN, k-NN, SVM, and ANFIS, is employed to the test data set for analysis. The assessments of their performances are done based on measures like RMSE, accuracy (i.e., classification accuracy), and kappa statistic as given in Table II.

Fig. 4. Image of statlog landsat satellite database.

TABLE II
PERFORMANCE COMPARISON USING LANDSAT SATELLITE DATABASE

Classifier	Classification Accuracy (%)	RMSE	Kappa statistic
NFS	**97.6**	**0.1446**	**0.9283**
RBFN	86.7	0.2798	0.8167
k-NN	87.5	0.2634	0.8243
SVM	85.4	0.2917	0.8051
ANFIS	90.7	0.2386	0.8651

From Table II, we can see that the NFS classification model has a classification accuracy of 97.6%. The accuracy values of RBFN, k-NN, SVM, and ANFIS models are 86.7%, 87.5%, 85.4%, and 90.7%, respectively. Thus, based on accuracy, NFS has performed better than RBFN, k-NN, SVM, and ANFIS. Then we analyze the performance of each classifier using RMSE and the kappa statistic values. The RMSE and kappa statistic index values of any classifier should lie between 0.0 and 1.0.

A lower RMSE value indicates a better classifier performance. The kappa statistic index estimates the accuracy for distinguishing between the reliability of the classified data gathered and their validity. We typically expect higher numeric value of kappa statistic for a specific classifier. We can see that the value of kappa statistic for the designated algorithms is around 0.81-1.0. Using the description of kappa statistic, the performances of these classification approaches indicate almost perfect agreement. Based on the result, NFS holds the first position with an RMSE measure of 0.1446 and a kappa measure of 0.9283. ANFIS does legitimately better with RMSE index as 0.2386 and kappa value as 0.8651. Followed by the k-NN model is having an RMSE measure of 0.2634 and a kappa measure of 0.8243, and then comes RBFN with an RMSE magnitude of 0.2798 and a kappa index of 0.8167. The SVM model is the worst performer with the highest RMSE measure (0.2917) and the lowest kappa measure (0.8051). Therefore, using evaluation measures such as RMSE, accuracy, and kappa statistic index, our proposed NFS classifier has accomplished the best.

Then, the classification models are compared using TP rate recall, FP rate, precision, and F-measure metrics resulting

TABLE III
DETAILED ACCURACY USING LANDSAT SATELLITE DATABASE

Classifier	TP-Rate /Recall	FP-Rate	Precision	F-Measure	AUC
NFS	**97.6 %**	**3.1 %**	**97.4 %**	**97.5 %**	**0.961**
RBFN	86.7 %	12.4 %	86.6 %	86.6 %	0.855
k-NN	87.5 %	10.7 %	87.4 %	87.4 %	0.869
SVM	85.3 %	13.5 %	85.2 %	85.2 %	0.847
ANFIS	90.6%	9.4 %	90.6 %	90.6 %	0.896

from the confusion matrix of individual classifier. The detailed accuracy for these classification models is presented using the percentage values (%) in Table III. We also compare the area under the receiver operating characteristic (ROC) curve (AUC) for different classifiers. The ROC curve typically plots *TP rate(t)* versus *FP rate(t)* with t as the changing parameter. For better classification performance, the AUC value of a classifier should come closer to one. Thus, models with greater AUC values are preferred to those with lesser AUC values. For assessing the performance of a classifier, we should assume higher magnitudes for TP rate/recall, precision, F-measure, and AUC, and a smaller magnitude for the FP rate.

From Table III, we can note that numeric values of TP rate/recall, FP rate, precision, and F-measure metrics for the NFS classifier are 97.6%, 3.1%, 97.4%, and 97.5% correspondingly. Whereas, the RBFN classifier is having these values as 86.7%, 12.4%, 86.6%, and 86.6%, respectively. The k-NN model has the values as 87.5%, 10.7%, 87.4%, and 87.4% separately. The SVM model has the values of TP rate/recall, FP rate, precision, and F-measure as 85.3%, 13.5%, 85.2%, and 85.2% individually. The ANFIS classifier is having these values as 90.6%, 9.4%, 90.6%, and 90.6%, respectively. Assuredly, NFS has the uppermost magnitudes for TP rate/recall, precision, and F-measure metrics and the lowermost magnitude for the FP rate compared with the other classifiers. The researchers typically consider F-measure as the greatest measure derived from the confusion matrix. Accordingly, the NFS model has the uppermost value for F-measure as 97.5%, the ANFIS model is having an F-measure value of 90.6%, and the RBFN classifier has the value for the F-measure as 86.6%. The k-NN model has got the value of F-measure as 87.4%. The SVM model has the value of F-measure as 85.2%.

Here, NFS has achieved the highest value for AUC (i.e., the value is 0.961) followed by the other models such as ANFIS, k-NN, RBFN, and SVM successively.

Undeniably, NFS classifier has performed significantly better than other classifiers in all respects.

2) Forest Covertype Database: Afterward, we use the UCI Forest Covertype [37] database to predict the soil classes constituting dissimilar forest cover types using cartographic variables. The current database was created choosing only a small segment (30×30 m^2 cell) from Region 2 of the United States Forest Service (USFS) Resource Information System data. The independent cartographic variables were resulting from records originally obtained from the USFS and the United States Geological Survey archives. The records were in raw form and consist of binary columns of information for independent qualitative variables representing wilderness areas and land types. This database was not developed using remotely sensed imagery data. Nevertheless, it is certainly a classical example of massive imagery used within the research community.

Fig. 5. Image of forest covertype database.

The area of study included four wilderness areas located in the Roosevelt National Forest in north-central Colorado of United States. They were, namely, Rawah (area 1), Neota (area 2), Comanche Peak (area 3), and Cache la Poudre (area 4). These zones represented forests with minimal human-induced disorders so that the existing forest cover types were more a consequence of ecological processes rather than management practices followed in forests. Of them, area 2 had the highest mean elevational value, followed by areas 1 and 3, while area 4 had the lowest mean elevational value.

This multivariate data set contains 581 012 rows and 55 columns. Of them, 54 columns of data (ten quantitative variables, four binary wilderness areas, and 40 binary soil type variables) are the input features and the last column denotes a class having seven soil class values. These soil class values are ranging from soil type 1 to soil type 7. As the database is huge in size, we use only 25% of records for simulation experiment. The approach should certainly build the models in lesser time compared with considering the full database. Here, this paper takes the full attribute list of the database for simulation. The image of the Covertype database is shown in Fig. 5.

Table IV presents the information related to the Forest Covertype data set attributes.

The implication of each of the attributes in the Table IV is described here. The first attribute elevation denotes the height using meters unit; the second attribute aspect is the value of perspective expressed in the unit degrees azimuth, while the third column slope indicates the slope value in degrees. The fourth and fifth columns denote the horizontal and vertical distances in meters to the nearest surface water features, respectively; while the sixth column signifies the distance to the nearest roadway using the same unit. The attribute numbers 7–9 indicate the hill shade index during the summer solstice at 9 A.M., at 12 noon, and at 3 P.M., respectively. The column number 10 denotes the horizontal distance to the nearest wildfire ignition points in meters.

TABLE IV
ATTRIBUTE INFORMATION OF COVERTYPE DATA SET

Sl. No.	Attribute	Data Type
1	Elevation	Quantitative
2	Aspect	Quantitative
3	Slope	Quantitative
4	Horizontal_Distance_To_Hydrology	Quantitative
5	Vertical_Distance_To_Hydrology	Quantitative
6	Horizontal_Distance_To_Roadways	Quantitative
7	Hillshade_9am	Quantitative
8	Hillshade_noon	Quantitative
9	Hillshade_3pm	Quantitative
10	Horizontal_Distance_To_Fire_Points	Quantitative
11-14	Wilderness_Area	Quantitative
15-54	Soil_Type	Quantitative
55	Cover_Type	Integer (1-7)

TABLE V
PERFORMANCE COMPARISON USING FOREST COVERTYPE DATABASE

Classifier	Classification Accuracy (%)	RMSE	Kappa statistic
NFS	**88.4**	**0.1806**	**0.7932**
RBFN	75.7	0.3192	0.6948
k-NN	76.8	0.2928	0.7039
SVM	78.3	0.2734	0.7231
ANFIS	80.3	0.2541	0.7521

The attribute numbers 11–14 represent four binary wilderness areas, while column numbers 15–54 designate 40 binary soil type variables. All the input attributes are quantitative, and their values should lie in the range of 0–255. The class attribute named Cover_Type has seven soil class types denoted by the integer numbers 1–7.

As usual, the five models, namely, NFS, RBFN, k-NN, SVM, and ANFIS, are employed to the test set for classification. We evaluate the performance of these classifiers along the base of different measures like accuracy, RMSE, and kappa statistic index as presented in Table V.

From Table V, we can see that NFS has a classification accuracy of 88.4%. The accuracy values of RBFN, k-NN, SVM, and ANFIS models are 75.7%, 76.8%, 78.3%, and 80.3%, respectively. Assuredly, accuracy wise NFS has performed much better than RBFN, k-NN, SVM, and ANFIS. Then we analyze the each classifier performance using RMSE and the kappa statistic measures. The kappa statistic for the designated algorithms is around 0.61–0.80. Using the concept of kappa statistic, the performance of these classification procedures is "substantial." Based on the result, NFS holds the first position with an RMSE measure of 0.1806 and a kappa statistic measure of 0.7932. ANFIS holds the succeeding position with RMSE value as 0.2541 and kappa index as 0.7521. Followed by SVM is having an RMSE measure of 0.2734 and a kappa measure of 0.7231, and then comes k-NN model with an RMSE magnitude of 0.2928 and a kappa index of 0.7039. RBFN stands last with the highest RMSE value being 0.3192 and the lowest kappa statistic value being 0.6948. Therefore, regarding the measures such as RMSE, accuracy, and kappa statistic index, NFS classifier has accomplished the finest.

TABLE VI
DETAILED ACCURACY USING FOREST COVERTYPE DATABASE

Classifier	TP-Rate /Recall	FP-Rate	Precision	F-Measure	AUC
NFS	**88.3 %**	**11.7 %**	**88.3 %**	**88.3 %**	**0.881**
RBFN	75.7 %	24.3 %	75.7 %	75.7 %	0.754
k-NN	76.6 %	23.4 %	76.6 %	76.6 %	0.765
SVM	78.3 %	21.7 %	78.3 %	78.3 %	0.782
ANFIS	80.2 %	19.8 %	80.2 %	80.2 %	0.801

Afterward, these models are compared for performance analysis using metrics such as TP rate/recall, FP rate, precision, F-measure, and AUC. Table VI provides the in-depth accurateness of these models.

From Table VI, we can realize that the values of TP rate/recall, FP rate, precision, F-measure, and AUC for NFS model are 88.3%, 11.7%, 88.3%, and 88.3% individually. Though, the RBFN classifier is having these values as 75.7%, 24.3%, 75.7%, and 75.7%, respectively. The k-NN model has the values as 76.6%, 23.4%, 76.5%, and 76.5% separately. The SVM model has the values of TP rate/recall, FP rate, precision, and F-measure as 78.3%, 21.7%, 78.3%, and 78.3% correspondingly. The ANFIS classifier is having these values as 80.2%, 19.8%, 80.2%, and 80.2%, respectively. NFS has the highest values for TP rate/recall, precision, and F-measure and the lowest value for FP rate among all. The NFS model has the highest magnitude for the F-measure as 88.3%. The ANFIS model is having an F-measure value of 80.2%; the RBFN classifier has the value for the F-measure as 75.6%, while the k-NN model has the value as 76.5%. The SVM model has got the value of F-measure as 78.3%. NFS has also attained the maximum value for AUC (i.e., the value is 0.881). Indeed, NFS has given an improved performance even with a lesser amount of data compared with other classifiers.

3) Wilt Database: Finally, we use the UCI Wilt [38] data set for classifying soil type related to the diseased trees. It is a high-resolution remotely sensed data set that contains some training and testing instances from a remote sensing survey in the Quickbird imagery. The multivariate imagery consists of 4889 tuples and has six attributes including the class attribute. The class attribute has two values, namely, w denoting diseased trees and class n indicating all other land cover. There are 74 instances of the w class and 4265 for n class in this data set. The data set consists of image

TABLE VII
ATTRIBUTE INFORMATION OF WILT DATA SET

Sl. No.	Attribute	Data Type
1	GLCM_Pan	Numeric
2	Mean_G	Numeric
3	Mean_R	Numeric
4	Mean_NIR	Numeric
5	SD_Pan	Numeric
6	Class	Categorical: 'w' (diseased trees), 'n' (all other land cover)

TABLE VIII
PERFORMANCE COMPARISON USING WILT DATABASE

Classifier	Classification Accuracy (%)	RMSE	Kappa statistic
NFS	98.3	0.1014	0.8732
RBFN	87.4	0.2048	0.7679
k-NN	88.2	0.1849	0.7836
SVM	86.5	0.2247	0.7457
ANFIS	90.2	0.1556	0.8274

segments, constructed by applying segmentation operation on the pan-sharpened image. These image segments consist of the spectral information resulting from the multispectral image bands of the Quickbird imagery. They also contain the texture information derived from the panchromatic image band (Pan Band). Table VII presents the attribute information of the Wilt data set.

The significance of each of the attributes in Table VII is identified here. The first attribute GLCM_Pan denotes the gray-level cooccurrence matrix with respect to the Pan band. The second, third, and fourth columns indicate the mean of green value, mean of red value, and mean of near infrared value, respectively. The fifth attribute SD_Pan implies the standard deviation concerning the panchromatic band. The database has a class attribute with two values, namely, w and n indicating diseased trees and all other land covers, respectively. Here, we consider the full attribute list of the database for simulation.

As usual, the five classification models are employed to the test set for performance analysis. We evaluate the operation of these classifiers using diverse measures like RMSE, accuracy, and the kappa statistic as presented in Table VIII.

From Table VIII, we can see that NFS model has the accuracy of 98.3%. The accuracy values of RBFN, k-NN, SVM, and ANFIS models are 87.4%, 88.2%, 86.5%, and 90.2%, respectively. Admittedly, NFS has performed better than RBFN, k-NN, SVM, and ANFIS in terms of accuracy. Then we analyze the each classifier performance using RMSE and the kappa statistic measures. The result shows that the kappa statistic measures of NFS and ANFIS models are around 0.81–1.0. These values of kappa statistic denote

TABLE IX
DETAILED ACCURACY USING WILT DATABASE

Classifier	TP-Rate /Recall	FP-Rate	Precision	F-Measure	AUC
NFS	**98.2 %**	**5.3 %**	**98.2 %**	**98.2 %**	**0.976**
RBFN	87.4 %	14.9 %	87.3 %	87.3 %	0.863
k-NN	88.2 %	12.7 %	88.1 %	88.1 %	0.875
SVM	86.5 %	16.8 %	86.4 %	86.4 %	0.859
ANFIS	90.1 %	9.9 %	90.1 %	90.1 %	0.897

"almost perfect agreement." While the kappa values of the RBFN, k-NN, and SVM methods are around 0.61–0.80, which denotes "substantial." According to the result, NFS holds the first position with an RMSE measure of 0.1014 and a kappa measure of 0.8732. ANFIS holds the next position with RMSE value as 0.1556 and a kappa measure as 0.8274. Followed by k-NN has an RMSE measure of 0.1849 and a kappa measure of 0.7836; and then comes the RBFN model with an RMSE magnitude of 0.2048 and a kappa index of 0.7679. SVM turns out to be the worst performer with the highest RMSE magnitude (0.2247) and the lowest kappa index (0.7457). Therefore, with reference to the evaluation measures such as RMSE, accuracy, and kappa statistic, NFS method has performed the greatest.

Then, we compare their performances using TP rate/recall, FP rate, precision, F-measure, and AUC. As usual, we should assume higher magnitudes for TP rate/recall, precision, F-measure, and AUC, and a smaller magnitude for the FP rate. The detailed accuracy for these classifiers is presented in Table IX.

From Table IX, we note that the values of TP rate/recall, FP rate, precision, and F-measure metrics for the NFS model are 98.2%, 5.3%, 98.2%, and 98.2% separately. The RBFN classifier is having these values as 87.4%, 14.9%, 87.3%, and 87.3%, respectively. The k-NN model has the values as 88.2%, 12.7%, 88.1%, and 88.1% separately. The SVM model has values of TP rate/recall, FP rate, precision, and F-measure as 86.5%, 16.8%, 86.4%, and 86.4% correspondingly. The ANFIS classifier is having these values as 90.1%, 9.9%, 90.1%, and 90.1%, respectively. Certainly, the NFS classifier has achieved the maximum values of TP rate/recall, FP rate, precision, and F-measure and the minimum values of FP rate. Seeing F-measure as the supreme performance measure, NFS model has the value for the F-measure as 98.2%. The ANFIS model has an F-measure value of 90.1%; the k-NN classifier is having the value for the F-measure as 88.1%, while the RBFN model has the F-measure value as 87.3%. The SVM model has got the value of F-measure as 86.4%. Here, NFS has achieved the highest value for AUC (i.e., the value is 0.976). Again, NFS has achieved meaningfully superior to other classifiers in all respects.

Concerning the different performance measures used, we have got excellent results for our proposed NFS classifier compared with RBFN-, k-NN-, SVM-, and ANFIS-based models. In fact, the NFS classifier has the highest values for accuracy, kappa statistic index, TP rate/recall, precision, F-measure, and AUC and the lowest values for RMSE and FP rate. An algorithm having exceptional preciseness and reduced error rate will be considered effective as because it has the greater classification capability and predictive power in the soil data mining field. Indeed, the proposed NFS model has outperformed RBFN, k-NN, SVM, and ANFIS classifiers in terms of the evaluation measures used in simulation.

VI. Conclusion

This paper deals with the determination of soil classes from large imagery databases. In our work, we propose an NF method for soil classification and establish its efficiency successfully using three UCI data sets, namely, Statlog Landsat Satellite, Forest Covertype, and Wilt. The method utilizes and integrates the primary benefits of ANNs such as immense parallelism, adaptivity, robustness, and optimality with the imprecision and vagueness managing capability of fuzzy sets. Furthermore, the proposed classification model builds a membership matrix that offers information of the feature-wise degree of membership of a data pattern to all the classes instead of considering a specific class. The property successively delivers improved generalization ability.

As a conclusion, we have taken our objective to investigate and analyze the proposed NF classifier for soil classification and then compare its performance with RBFN, k-NN, SVM, and ANFIS using different evaluation measures. These measures are RMSE, kappa statistic, accuracy, FP rate, TP rate (or recall), precision, F-measure, and AUC. The most promising technique established on performance evaluation along the three UCI data sets is our proposed NFS. It has an accuracy of 97.6% using the Statlog Landsat Satellite data set, 88.4% using the Forest Covertype data set, and 98.3% using the Wilt data set. These values are surely better than that of MLP, k-NN, SVM, and ANFIS classifiers. The NFS model also has the lowest RMSE value and the highest F-measure, AUC, and kappa statistic values compared with the other classifiers. Indeed, it has performed significantly better than other dominant classifiers being used here. Therefore, we can conclude that our proposed NFS classifier has the potential to replace the traditional classification approaches for utilization in the applied soil data mining field.

It is also observed that the proposed NF classifier offers an enhanced performance even with less training data. In the case of large imagery databases, the performance of the NFS classifier is meaningfully higher than the other predominant classification methods used (accuracy is more than 7%–8%). Thus, its learning capability with a lesser percentage of training data makes it practically applicable to any large imagery soil database with a vast number of input features and classes. The NF classification method proposed here is robust and efficient. However, a more extensive investigation and assessment of the proposed NF method will be the objective of our future research.

References

[1] R. Agrawal, T. Imielinski, and A. Swami, "Database mining: A performance perspective," *IEEE Trans. Knowl. Data Eng.*, vol. 5, no. 6, pp. 914–925, Dec. 1993.

[2] M.-S. Chen, J. Han, and P. S. Yu, "Data mining: An overview from a database perspective," *IEEE Trans. Knowl. Data Eng.*, vol. 8, no. 6, pp. 866–883, Dec. 1996.

[3] P. Bhargavi and S. Jyothi, "Soil classification using GATree," *Int. J. Comput. Sci. Inf. Technol.*, vol. 2, no. 5, pp. 184–191, Oct. 2010.

[4] S. J. Cunningham and G. Holmes, "Developing innovative applications in agriculture using data mining," in *Proc. Southeast Asia Regional Comput. Confederation Conf.*, Singapore, 1999, pp. 25–29.

[5] V. Ramesh and K. Ramar, "Classification of agricultural land soils: A data mining approach," *Agricult. J.*, vol. 6, no. 3, pp. 82–86, 2011.

[6] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, 2nd ed. San Mateo, CA, USA: Morgan Kaufmann, 2005, pp. 285–378.

[7] R. Rojas, "The backpropagation algorithm," in *Neural Networks: A Systematic Introduction*. Berlin, Germany: Springer-Verlag, 1996, pp. 151–184.

[8] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 1998, pp. 253–277.

[9] N. K. Bose and P. Liang, *Neural Network Fundamentals With Graphs, Algorithms, and Applications*. New York, NY, USA: McGraw-Hill, 1996, pp. 119–209.

[10] P. J. Werbos, *The Roots of Backpropagation: From Ordered Derivatives to Neural Networks and Political Forecasting*. New York, NY, USA: Wiley, 1994, pp. 67–163.

[11] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, Oct. 1986.

[12] K. J. Hunt, R. Haas, and R. Murray-Smith, "Extending the functional equivalence of radial basis function networks and fuzzy inference systems," *IEEE Trans. Neural Netw.*, vol. 7, no. 3, pp. 776–781, May 1996.

[13] D. S. Broomhead and D. Lowe, "Radial basis functions, multi-variable functional interpolation and adaptive networks," Roy. Signals Radar Establishment, New York, NY, USA, Tech. Rep. 4148, 1988, pp. 1–34.

[14] G. Shakhnarovich, T. Darrell, and P. Indyk, Eds., *Nearest-Neighbor Methods in Learning and Vision: Theory and Practice*. Cambridge, MA, USA: MIT Press, 2005.

[15] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995.

[16] L. A. Zadeh, "Fuzzy sets," *Inf. Control*, vol. 8, no. 3, pp. 338–353, Jun. 1965.

[17] B. Liu, *Uncertainty Theory: An Introduction to Its Axiomatic Foundations*. Berlin, Germany: Springer-Verlag, 2004, pp. 191–346.

[18] D. Dubois and H. M. Prade, *Fuzzy Sets and Systems: Theory and Applications*. New York, NY, USA: Academic, 1980, pp. 255–348.

[19] J.-S. R. Jang, C.-T. Sun, and E. Mizutani, *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1997, pp. 333–393.

[20] C.-T. Lin and C. S. Lee, *Neural-Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1996, pp. 313–459.

[21] D. Nauck, F. Klawonn, and R. Kruse, *Foundations of Neuro-Fuzzy Systems*. Chichester, U.K.: Wiley, 1997.

[22] J. A. Shine and D. B. Carr, "Relationships between land cover and spatial statistical compression capabilities in high-resolution imagery," in *Proc. 34th Interface Symp.*, Montreal, QC, Canada, Apr. 2002, p. 450.

[23] J. A. Shine, "Mapping and modelling 1-meter multispectral imagery data," in *Proc. Amer. Statist. Assoc.*, Alexandria, VA, USA, 2000, pp. 81–83.

[24] J. A. Shine, "Compression and analysis of very large imagery data sets using spatial statistics," in *Proc. 33rd Symp. Interface*, Costa Mesa, CA, USA, Jun. 2001, pp. 1–6.

[25] J. A. Shine and D. B. Carr, "A comparison of classification methods for large imagery data sets," in *Proc. Joint Statist. Meetings Statist. ERA Technol. Change-Statist. Comput. Section*, New York, NY, USA, Aug. 2002, pp. 3205–3207.

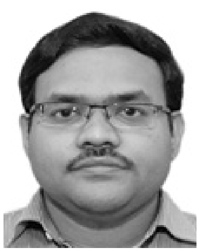
[26] O. Rozenstein and A. Karnieli, "Comparison of methods for land-use classification incorporating remote sensing and GIS inputs," *EARSeL eProc.*, vol. 10, no. 1, pp. 27–45, 2011.

[27] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *Int. J. Remote Sens.*, vol. 28, no. 5, pp. 823–870, Mar. 2007.

[28] N. Alajlan, Y. Bazi, F. Melgani, and R. R. Yager, "Fusion of supervised and unsupervised learning for improved classification of hyperspectral images," *Inf. Sci.*, vol. 217, pp. 39–55, Dec. 2012.

[29] M. Li, S. Zang, B. Zhang, S. Li, and C. Wu, "A review of remote sensing image classification techniques: The role of spatio-contextual information," *Eur. J. Remote Sens.*, vol. 47, pp. 389–411, Jun. 2014.

[30] S. Ghosh, S. Biswas, D. Sarkar, and P. P. Sarkar, "A novel neuro-fuzzy classification technique for data mining," *Egyptian Informat. J.*, vol. 15, no. 3, pp. 129–147, Nov. 2014.
[31] A. Elisseeff and H. Paugam-Moisy, "Size of multilayer networks for exact learning: Analytic approach," in *Advances in Neural Information Processing Systems*, vol. 9. Cambridge, MA, USA: MIT Press, Dec. 1996, pp. 162–168.
[32] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Stanford, CA, USA: Springer, 2009, pp. 241–249.
[33] J. S. Armstrong and F. Collopy, "Error measures for generalizing about forecasting methods: Empirical comparisons," *Int. J. Forecasting*, vol. 8, no. 1, pp. 69–80, Jun. 1992.
[34] J. Carletta, "Assessing agreement on classification tasks: The kappa statistic," *Comput. Linguistics*, vol. 22, no. 2, pp. 249–254, Jun. 1996.
[35] S. V. Stehman, "Selecting and interpreting measures of thematic classification accuracy," *Remote Sens. Environ.*, vol. 62, no. 1, pp. 77–89, Oct. 1997.
[36] *Statlog Landsat Satellite Data Set: Machine Learning Repository*, Univ. California, Irvine, Irvine, CA, USA, 1993.
[37] *Covertype Data Set: Machine Learning Repository*, Univ. California, Irvine, Irvine, CA, USA, 1998.
[38] *Wilt Data Set: Machine Learning Repository*, Univ. California, Irvine, Irvine, CA, USA, 2014.



Soumadip Ghosh received the B.Tech. and M.Tech. degrees from Kalyani University and Calcutta University in 2002 and 2005, respectively. He received the Ph.D. degree from Kalyani University.

Currently, he acts as an assistant professor with the Academy of Technology, India. His research interests include machine learning, data mining, Internet technology, cryptography, and network security.

Debasish Biswas received the M.Sc. and M.Tech. degrees from Calcutta University, in 2003 and 2005, respectively. He is now a Ph.D. candidate at Kalyani University.

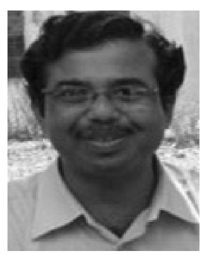
Currently, he acts as an Assistant Professor with the Dr. Meghnad Saha Institute of Technology, Haldia. His research interests include image processing, steganography, cryptography, and network security.

Sushanta Biswas was felicitated with the Ph.D. degree in engineering from Jadavpur University in 2004.

He is currently working as an Associate Professor with the Department of Engineering and Technological Studies, University of Kalyani. His area of research includes machine learning, microstrip antenna, microstrip filter, frequency selective surfaces, and artificial neural network.

Debasree Chanda (Sarkar) was felicitated with the Ph.D. degree in engineering from Jadavpur University, in 2004.

She is currently working as an Associate Professor with the Department of Engineering and Technological Studies, University of Kalyani. Her area of research includes microstrip antenna, microstrip filter, frequency selective surfaces, and artificial neural network.



Partha Pratim Sarkar was felicitated with the Ph.D. degree in engineering from Jadavpur University, in 2002.

He is currently working as a Professor with the Department of Engineering and Technological Studies, University of Kalyani. His area of research includes microstrip antenna, microstrip filter, frequency selective surfaces, and artificial neural network.

Dr. Sarkar is a Life Fellow of IETE and IE (India).