Research Article

Evaluation and comparison of eight machine learning models in land use/land cover mapping using Landsat 8 OLI: a case study of the northern region of Iran

Ali Jamali[1]

Abstract

Land use land cover change mapping has been used for monitoring environmental changes as an essential factor to study on the earth's surface land cover in the field of climate change phenomena such as floods and droughts. Remote sensing images have been suggested to present inexpensive and fine-scale data offering multi-temporal coverage. This tool is useful in the field of environmental monitoring, land-cover mapping, and urban planning. This study aims to evaluate eight machine learning algorithms for image classification implemented in WEKA and R programming language. Firstly, Landsat 8 OLI/TIRS Level-2 images based on eight machine learning techniques including Random Forest, Decision Table, DTNB, J48, Lazy IBK, Multilayer Perceptron, Non-Nested Generalized Exemplars (NN ge), and Simple Logistic are classified. Then, obtained results are compared in term of Overall Accuracy (OA), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE) for land use land cover mapping. Among the eight machine learning algorithms used for image classification based on the training and test dataset, NN ge classifier is ranked first with values of 100, 0, and 0 for Overall Accuracy, Mean Absolute Error and Root Mean Squared Error respectively. All machine learning algorithms had an Overall Accuracy of more than 99% for the training dataset. On the other hand, for the test dataset, J48 and DTNB algorithms had the worst performance with values of 88.1188 and 76.9802 respectively for the Overall Accuracy.

Keywords Image classification · Machine learning · LULCC · R statistical packages · WEKA

1 Introduction

The changes in the climate of the earth such as the increase in the daily earth surface temperature and the need for monitoring its impacts on the earth's surface call for environmental monitoring approaches. Land cover is known as one of the fundamental terrestrial climate variables [1, 2]. In the land cover mapping, detailed land cover maps are an essential input for various scientific associations working on climate change investigations, sustainable development, geomorphology, and social knowledge management of natural resources, and monitoring the agricultural lands. Meanwhile, an unprecedented volume of satellite imagery information along with an enhanced

level of spatial and spectral as well as temporal resolution is provided by open data policies of the USA and EU countries. The capacity of land cover monitoring is currently enhanced via utilizing dense time series, demanding efficient, cost-effective classification approaches which have become possible due to the availability of Landsat satellite imagery.

There is a definite need for precise and timely LULC information to monitor and analyze the human and physical environment. These data can be utilized to a multitude of distinct domains (e.g., mobility to health, ecology, agriculture, risk analyzing, and policies of management) as a consolidated practice [3]. Currently, for fine-scale analyzing of the earth's surface, the imagery of Medium Spatial

Resolution (MSR) are used in remote sensing. In order to confirm decisions in different domains, such as agriculture, biodiversity, and also environmental prediction, the MSR land cover classification can be performed. Land cover is known as an essential factor which may link and influence various fields namely human and physical environment [4].

Land cover change is a significant parameter leading to global change. Additionally, it can affect ecological approaches [5] as well as the earth's conditions, both of which are related to climatic change [6]. Earth observation satellites sensor data is known as an effective factor in investigating the results of climate change. Land cover mapping are among the most important application of earth observation satellites sensor data. Change of the land cover may affect the climate through manipulation of the composition of pollutant emissions like carbon dioxide [7, 8]. Today, LULC statistics are the prerequisites for policy and decision making strategies which have an effect on societies and their economies [9].

There are different classification methods from unsupervised algorithms including K-means clustering, parametric algorithms such as maximum likelihood [10], machine learning algorithms including Artificial Neural Networks (ANNs) and SVMs [11, 12], decision trees [13, 14], and ensemble of classifiers [15]. Algorithms of machine learning commonly have high accuracy and efficiency in comparison to usual parametric algorithms for dealing with large and assembled databases [16]. Land cover classification is usually used to create a thematic map of the land cover. It consists of the features such as water, soil, vegetation, and man-made structures [17]. The number of classes in the image can be decided based on sensor resolutions. Supervised classification approaches have been known to be superior compared to the unsupervised ones in the field of land cover mapping; however, precise and sufficient training information should be used. In various studies, machine learning algorithms (e.g., Support Vector Machines [18, 19], Random Forests [20], and ANNs have been considered quite influential for classification purposes. In the classification process of land cover, the SVM is widely used [12, 21]. However, several researches indicate that RF and ANN algorithms have been outperformed by other algorithms [22, 23]. SVM provides the capacity to achieve high classification precisions even with small training data [24]. In addition, the SVM is a robust approach for low noise levels, even in the attendance of mislabeled training information [22].

Classification of the image is presented as an image processing method which defines features in each image based on their spectral signatures, which is considered the reflection of a wavelength function. Each feature such as water region has a specific signature that will be used in feature classification [25]. In the case of environmental studies, images of Landsat have been widely used. Several researches have created land cover maps by the supervised classification methods using Landsat satellite imagery. Land cover maps, originally, can be used to present time-series changes in urban development and green areas. Additionally, the relationship between land cover change and urban population has been investigated by a number of researchers [26–33]. Objectives of this research are defined as:

1. Implementation of eight advanced machine learning algorithms including Random Forest, Decision Table, DTNB, J48, Lazy IBK, Multilayer Perceptron, NN ge, and Simple Logistic for image classification within WEKA and R programming language.
2. Land use land cover mapping of the northern region of Iran based on the open data source of Landsat 8 OLI imagery.
3. Evaluation of the eight advanced machine learning algorithms for image classification in term of the OA, MAE, and RMSE.
4. Proposing a fit-for-purpose machine learning algorithm for Northern Iran for LULCC mapping.

2 Study area and data collection

Figure 1 presents the information obtained by the Landsat 8 OLI satellite (for 21st February 2019 data collection of the City of Sari in WGS 84/UTM area 39 N). Sari (study area) is located in the north of Iran, between the southern coast of the Caspian Sea and the northern slopes of the Alborz Mountains and it is the largest and most populous city of Mazandaran Province. Sari is the capital of and the largest and most populous city of the province of Mazandaran. It has a humid subtropical climate with a Mediterranean climate influence. Winters are cold and rainy while summers are hot and humid. In this research, for image classification, Sari was chosen because of the complexity of nature composition. Mazandaran with the Caspian Hyrcanian mixed forests ecoregion has a diverse nature including plains, forest, rainforest, and prairies from snowcapped alborz to sand beaches of the Caspian see. Mazandran has a population of more than three million with a density of $130/km^2$. In Mazandran province, the Dohezar, Sehezar, and Kojoor forest watersheds are located.

For image classification, Landsat 8 OLI bands with wavelength from 0.4430 to 2.2010 μm including coastal aerosol (0.4430), Blue (0.4826), Green (0.5613), Red (0.6546), Near Infrared (0.8646), Shortwave Infrared 1 (1.6090), and Shortwave Infrared 2 (2.2010) are used.

Fig. 1 The study area: **a** map of Iran **b** Mazandaran province **c** image of the study area
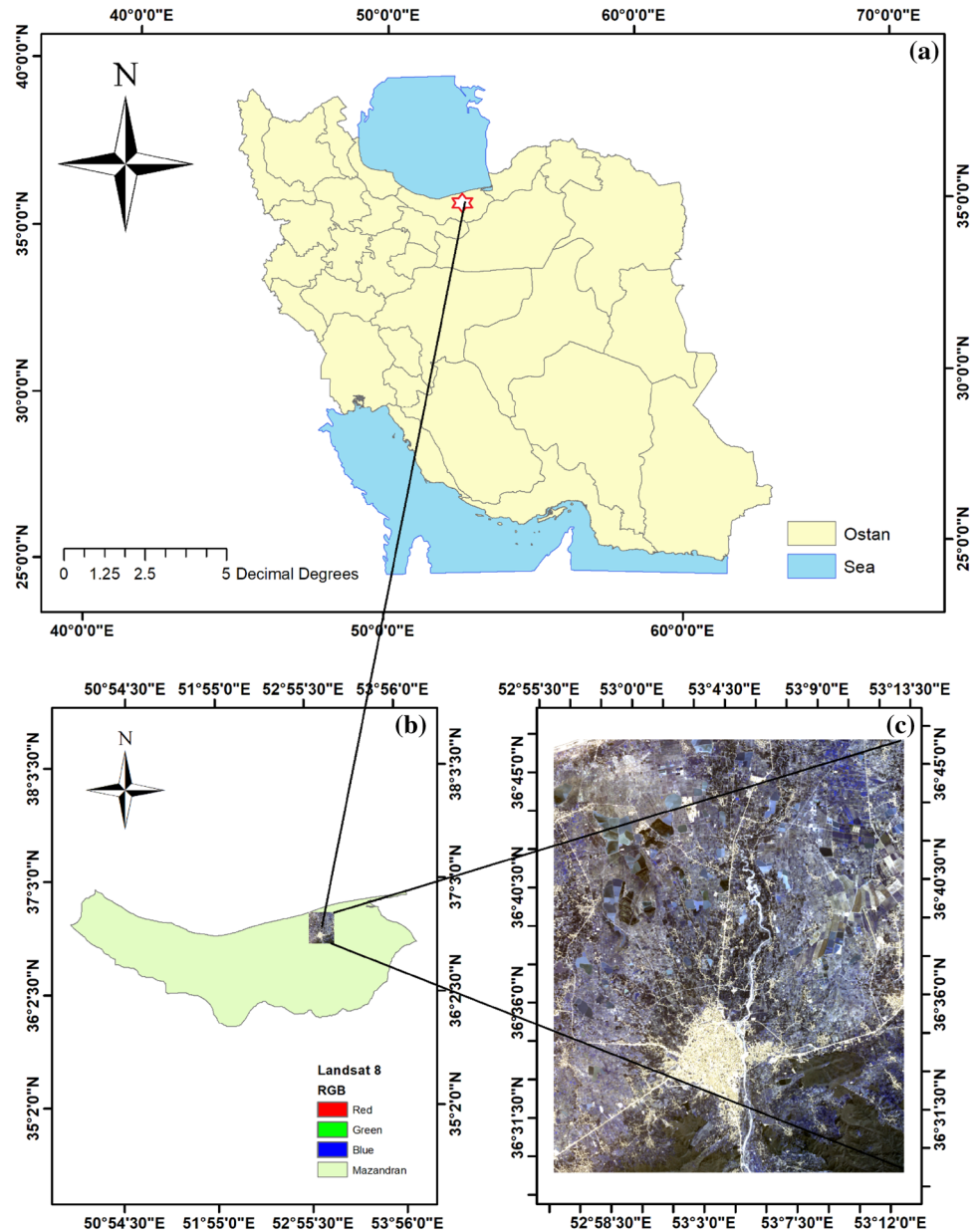
3 Methodology

Figure 2 presents the methodology of this research. The Landsat 8 OLI/TIRS Level-2 image was corrected from the radiometric and atmospheric effects in the first step. Second, the reference data, including the training and testing samples, were carefully generated so as to represent four LULC classes as follows: built-up areas, bare soil, vegetation, and roads. Next, various machine learning algorithms were employed to classify the image of the study area in the third step. Fourth, the outputs of the best predictive models were statistically assessed. Finally, the results of different image classification algorithms were discussed.

3.1 Pre-processing

In the present study, an atmospheric correction is done because various atmospheric evaluations are required in order to predict the reflectance to the ground (ρ) within the pre-processing stage of the pictures. We used Dark Object Subtraction (DOS) among image-based atmospheric correction approaches for atmospheric correction. For satellites images classification task, various levels of classes of mean built-in, wetlands, and crop are predicted using a composition of different spectral signatures of district bands with or without Normalized Vegetation Difference Index (NDVI) (as can be observed in Eq. 1 [34]):

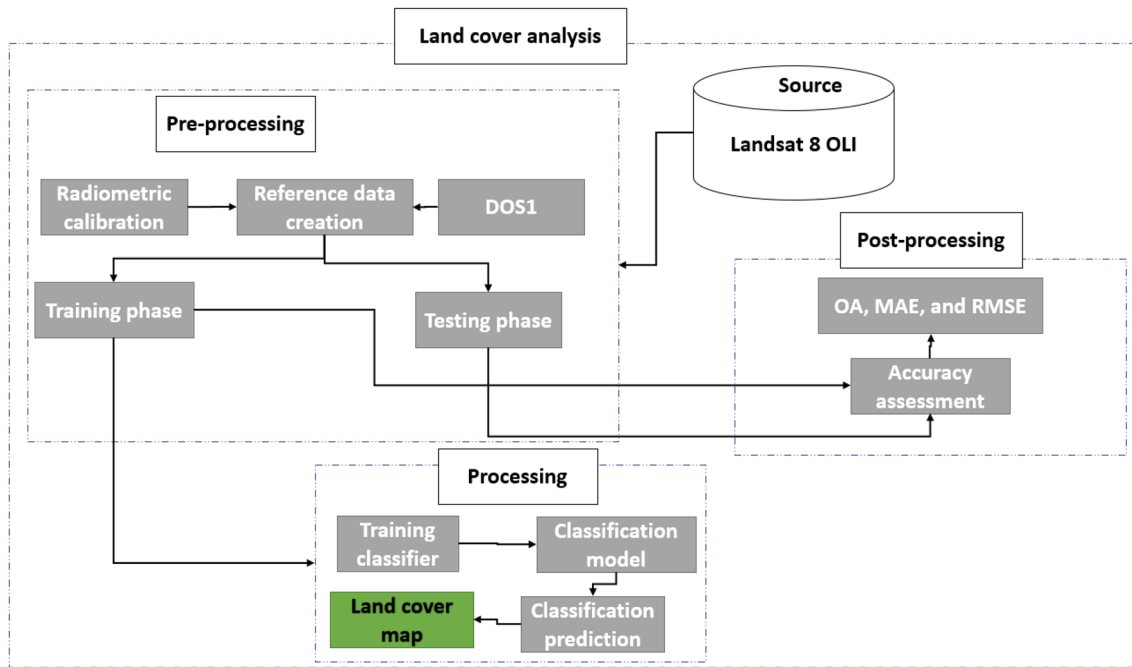$$NDVI = (NIR - R)/(NIR + R) \tag{1}$$

Fig. 2 The workflow of this research

where *NIR* stands for Near Infra-Red band and *R* stands for Red band.

Spectral indices are highly regarded as the basic collection of input characteristics in the case of land cover classification [35]. According to similar researches [36, 37], the seven atmospherically reformed L8 OLI/TIRS spectral bands are utilized. To improve LULC classification, the enhanced vegetation index and the normalized difference buildup index were also implemented in the analysis (Eqs. 2, 3):

$$NDBI = \frac{B - SWIR}{B + SWIR} \qquad (2)$$

$$EVI = 2.5 * \frac{NIR - R}{(NIR + 6R - 7.5B) + 1} \qquad (3)$$

3.2 Classification

Eight advanced machine learning algorithms including Random Forest, Decision Table, DTNB, J48, Lazy IBK, Multilayer Perceptron, NN ge, and Simple Logistic were applied for image classification in the study area.

3.2.1 Random Forest (RF)

The approach of Random Forest (proposed by Breiman [38]) is obtained by developing the classification and also regression trees-CART [13]. The method of RF is an ensemble learning method commonly used in the case of land-cover classification by taking advantage of multispectral and hyperspectral satellite sensor imagery. The RF produces various trees according to random bootstrapped of the training database patterns. This method performs random binary trees which produce a training subset above bootstrapping approach. In addition, a random choice of the training information is employed and accomplished to create the model from the initial database; however, out-of-the bag (OOB) is known as the data that is not involved [39]. A Random Forest is defined as a predictor consisting of a collection of randomized base regression trees $\{r_n(x, \theta_n, D_n), m \geq 1\}$.. These random trees are integrated to create the aggregated regression estimate as $\bar{r}_n(X, D_n) = E_\theta[r_n(X, \theta, D_n)]$, where E_θ defines expectation regarding random parameters, conditionally on X and dataset D_n.

3.2.2 Decision Table

Decision Table (DT) is a classifier which uses a simple DT majority classifier. DTs are among the simplest hypothesis spaces possible, and usually, easy-to-use [40]. In addition, it has two parts, including a set of characteristics involved in the table along with a body including labelled samples of the space specified using the features. A DT arranger finds exact accommodations in the DT by utilizing only the properties in the schema with the consideration of an

unlabeled instance. However, it is crucial to note that there can be other matching samples in the table.

3.2.3 DTNB

In order to construct and utilize a decision table or naive Bayes hybrid arranger, DTNB as an appropriate classifier can be employed [41]. In the present study, our employed algorithms analyse the merit of separating the properties into two disintegrate subsets of the decision table and also the naive Bayes. Additionally, whole properties were initially modelled using the decision table. The used algorithm also attends dropping a property from the model.

3.2.4 J48 algorithm

The algorithm of J48 is known as a suitable classifier which creates an unpruned and pruned decision tree of C4.5 [42]. The trees generated by C4.5 usually are small and exact. Moreover, using a decision tree of C4.5 leads to fast and reliable classifications for various domains. These appropriate characteristics of C4.5 approach make decision trees (such as the C4.5) a noteworthy and popular tool in the case of classification tasks. Researchers have broadly suggested decision trees due to their benefits.

3.2.5 Lazy IBK

Lazy IBK as a K-nearest neighbour classifier can choose a proper value of K according to cross-validation [43]. This algorithm is known as the most popular algorithm used for pattern recognition. The algorithm of KNN is a method of

Lazy learning when the function is just predicted locally, and also all calculations are deferred up to classification. In this way, an object can be classified using a majority of its neighbours. In this algorithm, K stands to a positive integer. The neighbours commonly have been chosen from a collection of objects in WEKA approach named IBK.

3.2.6 Multilayer Perceptron

Deep learning approach is highly developed in different tasks such as image classification, object recognition and also semantic comprehension of natural pictures. Convolutional neural networks (CNNs) are widely applied to categorize remote sensing images. Artificial neural networks are commonly designed by taking advantage of various interconnected nodes (e. i., neurons). In an artificial neural network (ANN), there may be three layers involving the input layer as well as the hidden and output layer. Information needs to be classified into three databases as training, validation also exam obtained data in order to train the ANN. There are different ways to specify the most appropriate number of hidden neurons. However, in this regard, training various networks is suggested as the most appropriate approach. The neural network of Multi-layer perceptron (MLP) is a widely utilized ANNs that identifies itself by using three layers [44]. These layers are commonly defined using layers namely input, hidden, and output and including various computational modules named nodes or neurons (see Fig. 3).

A perceptron creates a single output according to several real-valued inputs by generating a linear combination using its input weights as $y = \varphi \sum_{i=1}^{n} w_i x_i + b = \varphi(w^T x + b),$
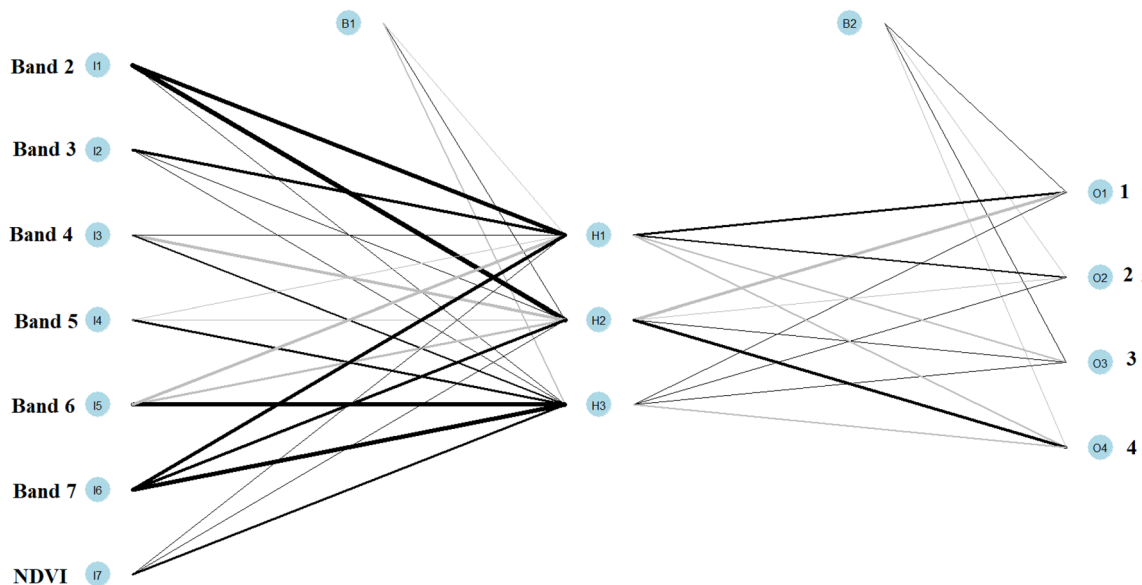

Fig. 3 The artificial neural network architecture for image classification

where w is vector of weights, x is vector of inputs, b is the bias, and φ is the non-linear activation function.

3.2.7 NN ge

The NN ge accomplishes generalization task utilizing merging samples. However, it forms hyper-rectangles in property space, which demonstrates conjunctive rules together with internal disjunction. By connecting this algorithm to its nearest neighbour of the similar class, the algorithm creates an extension each time a different sample is added to the data set, by joining it to its nearest neighbour of the same class. Firstly, NN ge learns incrementally using first categorizing. Next, it generalizes each new sample. One or even more hyper-rectangles can be determined which the new example is a member. The algorithm of NN ge simplifies these and as a result, the new instance is not a considered member. After classification tasks are carried out, the new sample, combined with the nearest model of the same class as an example or a hyper-rectangle. NN ge consists of an algorithm of Nearest-neighbor-like utilizing non-nested generalized samples in the form of hyperrectangles and considered as if–then rules [45].

3.2.8 Simple Logistic

In the simple Logistic algorithm, LogitBoost together with simple regression functions as basis learners were utilized in order to fit the logistic models [46]. This algorithm was classified in the group learning methods that used additive logistic regression by utilizing instance regression functions as basis learners. This algorithm also finds a function which can appropriately fit the training information using calculating the weights which amplify the log-likelihood function of the logistic regression. A logistic function is defined as $f(x) = \frac{L}{1+e^{-k(x-x_0)}}$ where e is Euler's number, x_0 is the x-value of the sigmoid's midpoint, L is maximum value of the curve, and k is the logistic growth rate.

4 Accuracy assessment and validation

Image classification of Landsat 8 OLI/TIRS Level-2 based on the proposed algorithms of machine learning (Random Forest, Decision Table, DTNB, J48, Lazy IBK, Multilayer Perceptron, NN ge, and Simple Logistic) are evaluated using OA, MAE, and RMSE indices (Eqs. 4–6) [47].

$$OA = \frac{number\ of\ correctly\ classified\ values}{total\ number\ of\ reference\ values} \tag{4}$$

$$MAE = \frac{\sum_{i=1}^{n} \left| Observed_i - Predicted_i \right|}{n} \tag{5}$$

$$RMSE = \frac{\sum_{i=1}^{n} \left(Observed_i - Predicted_i \right)^2}{n} \tag{6}$$

The numbers of training and testing objects including the build-up, soil, water, and vegetation regions are presented in Table 1.

Using machine learning algorithms to prevent overfitting issue a tenfold cross-validation technique is used.

5 Results and discussion

Finally, the outputs of image classifications are compared in term of OA, MAE, and RMSE.

5.1 Image classification algorithms ranking

The results of evaluation of classification algorithms based on the training data set and testing data set are presented in Tables 2 and 3. The results of proposed methods (e.g., Multilayer Perceptron, Simple Logistic, J48, Lazy IBK, Random Forest, Decision Table, DTNB, NN ge) for both of the training and testing data sets are assessed based on their predictive network results. The model's assessment are performed based on the three well known statistical indices as discussed earlier.

For the ranking based on the OA, MAE, and RMSE, scores from 8 (best performance) to 1 (least performance) is given to each machine learning algorithms implemented in WEKA and R programming language wherein case that several algorithms show equal performance they receive an equal score. OA varies from 0 to 100 (100 means that the algorithm has predicted all the classes correctly where 0 stands for 100% incorrect classification) where the algorithms with the highest scores receive a value of 8. The second algorithm with the highest OA will receive a value of 7, and this ranking system continues for the eight machine learning algorithms. The algorithm with the worst performance will receive a value of 1. For example, if three algorithms receive a score of 8 due to their equal performance, the fourth algorithm will receive a score of 5. Keeping in

Table 1 The number of training and testing data for image classification

Classes	Training set size	Validation set size
Build-up	71	48
Soil	84	82
water	207	125
Vegetation	98	149

Table 2 Image classification algorithms evaluation based on the training data set. It should be noticed that OA for Multilayer Perceptron and Simple Logistic classifiers are 100% but their MAE and RMSE are not equal to zero due to their prediction probabilities which are not equal to 1 (prediction probabilities varies from 0 to 1)

	Proposed models	Network results			Score of the predicted models			Total ranking score	Rank
		OA	MAE	RMSE	OA	MAE	RMSE		
1	Multilayer Perceptron	100	0.0042	0.006	8	4	6	18	4
2	Simple Logistic	100	0.122	0.1713	8	1	1	10	5
3	J48	99.3478	0.0033	0.0571	2	6	2	10	5
4	Lazy IBK	100	0.0036	0.0041	8	5	7	20	2
5	Random Forest	100	0.0017	0.0183	8	7	5	20	2
6	Decision Table	99.5652	0.017	0.0492	3	2	3	8	7
7	DTNB	99.3478	0.0066	0.0478	2	3	4	9	8
8	NN ge	100	0	0	8	8	8	24	1

Table 3 The rank of image classification algorithms evaluation based on the test data set. It should be noticed that OA for Multilayer Perceptron, Lazy IBK and Random Forest classifiers is 100% but their MAE and RMSE are not equal to zero due to their prediction probabilities which are not equal to 1 (prediction probabilities varies from 0 to 1)

	Proposed models	Network results			Score of the predicted models			Total ranking score	Rank
		OA	MAE	RMSE	OA	MAE	RMSE		
1	Multilayer Perceptron	100	0.0042	0.0075	8	6	6	20	3
2	Simple Logistic	97.0297	0.1492	0.2039	4	1	3	8	6
3	J48	88.1188	0.0594	0.2437	2	3	2	7	7
4	Lazy IBK	100	0.0032	0.0037	8	7	7	22	2
5	Random Forest	100	0.0074	0.0476	8	5	5	18	4
6	Decision Table	95.297	0.0361	0.1512	3	4	4	11	5
7	DTNB	76.9802	0.1257	0.309	1	2	1	4	8
8	NN ge	100	0	0	8	8	8	24	1

mind that scores are independent of the value of OA. MAE and RMSE vary from 0 to 1 (0 means that there is no error for the prediction of classes and 1 stands for 100% incorrect classification). The algorithm with the least RMSE and MAE will receive a value of 8 where the algorithm with the highest RMSE and MAE will receive a value of 1. Finally, the algorithm with the highest values will be ranked first where the algorithm with the lowest values will be ranked as the eighth-ranked algorithm.

For the training data set, NN ge classifier with values of 100, 0, and 0 for OA, MAE and RMSE shows the best performance. Random Forest classifier with values of 100, 0.0017, and 0.0183 is ranked in second place. Lazy IBK classifier are ranked second as well with values of 100, 0.0036, and 0.0041. Multilayer Perceptron classifier with values of 100, 0.0042, and 0.006 is ranked fourth. J48 classifier is ranked fifth with values of 99.3478, 0.0033, and 0.0571. Simple Logistic classifier with values of 100, 0.122, and 0.1713 is ranked fifth as well. Decision Table classifier with values of 99.5652, 0.017, and 0.0492 claims the seventh rank. DTNB classifier with values of 99.3478, 0.0066, and 0.0478 is ranked eighth.

On the other hand, and based on the testing data set, eight advanced mathematical and machine learning algorithms used in this research including Random Forest, Decision Table, DTNB, J48, Lazy IBK, Multilayer Perceptron, NN ge, Simple Logistic were compared and ranked with respect to OA, MAE and RMSE (Table 3).

For the test data set, NN ge classifier with values of 100, 0, and 0 for OA, MAE and RMSE shows the best performance. Lazy IBK classifier is ranked second with values of 100, 0.0032, and 0.0037. Multilayer Perceptron classifier with values of 100, 0.0042 and 0.0075 is ranked third. Fourth place belongs to Random Forest classifier with values of 100, 0.0074, and 0.0476. Decision Table classifier with values of 95.297, 0.0361, and 0.1512 is ranked fifth. With values of 97.0297, 0.1492, and 0.2039, Simple Logistic classifier is ranked sixth. J48 classifier is ranked seventh with values of 88.1188, 0.0594, and 0.2437. DTNB classifier has the worst performance with values of 76.9802, 0.1257, and 0.309.

As seen in Table 4, based on the training and test data set, NN ge classifier has the best performance in terms of OA, MAE, and RMSE with a total value of 48. Lazy IBK classifier is ranked second with a total value of 42. The third-ranked classifiers are Multilayer Perceptron and Random Forest with a total value of 38. Decision Table classifier with a total value of 19 is ranked fifth. With a total value of 18, Simple Logistic classifier is ranked sixth. J48 classifier is ranked seventh with a total value of 17.

Table 4 Total ranking score and ranking of the proposed classification models based on both training and testing data sets

	Proposed models	Score of the predicted models						Total score	Rank
		Training data set			Testing data set				
		OA	MAE	RMSE	OA	MAE	RMSE		
1	Multilayer Perceptron	8	4	6	8	6	6	38	3
2	Simple Logistic	8	1	1	4	1	3	18	6
3	J48	2	6	2	2	3	2	17	7
4	Lazy IBK	8	5	7	8	7	7	42	2
5	Random Forest	8	7	5	8	5	5	38	3
6	Decision Table	3	2	3	3	4	4	19	5
7	DTNB	2	3	4	1	2	1	13	8
8	NN ge	8	8	8	8	8	8	48	1

The eighth-ranked classifier is DTNB classifier with a total value of 13.

Table 5 presents the confusion matrix of the proposed machine learning algorithms based on the test data set where misclassification values are seen.

Considering that small training and test datasets are used in this research, NNge classifier had a perfect performance compared to other discussed machine learning algorithms. All algorithms had excellent performance with an OA of more than 99% for the training dataset. For the test dataset, J48 and DTNB classifier with OA of 88.1188 and 76.9802 had the worst performance where others mentioned classifiers had an OA of more than 95%.

5.2 Land use land cover maps

To provide a reliable estimation of the environmental assessment using Landsat 8 OLI for the city of Sari, the outputs of eight different machine learning classification techniques are provided in this study. In this regard, image classification based on the eight advanced mathematical and machine learning algorithms including Random Forest, Decision Table, DTNB, J48, Lazy IBK, Multilayer Perceptron, NN ge, Simple Logistic are presented in Fig. 4. Images are classified by four materials including the build-up, water, soil, and vegetation regions. Out of 879,048 cells in the study area, the DT classifier classified 115,938 cells as Build-up regions, 115,678 cells as soil regions, 610,969 as vegetation regions, and 36,436 cells as water regions. The DTNB classifier classified 168,636 cells as Build-up regions, 47,195 cells as soil regions, 572,948 as vegetation regions, and 90,269 cells as water regions. Also, 168,686 cells were classified as Build-up regions, 91,322 cells as soil regions, 611,796 as vegetation regions, and 15,244 cells as water regions by J48 classifier. 144,272 cells were classified as Build-up regions, 80,418 cells as soil regions, 607,966 as vegetation regions and 46,392 cells as water regions by Multilayer Perceptron classifier. NN ge classifier classified 149,154 cells as Build-up regions,

Table 5 Confusion matrix of advanced machine learning algorithms

	Model	Build-up	Soil	Vegetation	Water
Decision Table	Build-up	47	1	0	0
	Soil	0	64	18	0
	Vegetation	0	0	149	0
	Water	0	0	0	125
Multilayer Perceptron	Build-up	48	0	0	0
	Soil	0	82	0	0
	Vegetation	0	0	149	0
	Water	0	0	0	125
Simple Logistic	Build-up	48	0	0	0
	Soil	0	72	10	0
	Vegetation	0	0	149	0
	Water	0	0	0	125
J48	Build-up	48	0	0	0
	Soil	0	64	18	0
	Vegetation	0	0	149	0
	Water	0	30	0	95
Lazy IBK	Build-up	48	0	0	0
	Soil	0	82	0	0
	Vegetation	0	0	149	0
	Water	0	0	0	125
Random Forest	Build-up	48	0	0	0
	Soil	0	82	0	0
	Vegetation	0	0	149	0
	Water	0	0	0	125
DTNB	Build-up	48	0	0	0
	Soil	0	64	18	0
	Vegetation	0	0	149	0
	Water	0	52	23	50
NN ge	Build-up	47	1	0	0
	Soil	0	64	18	0
	Vegetation	0	0	149	0
	Water	0	0	0	125
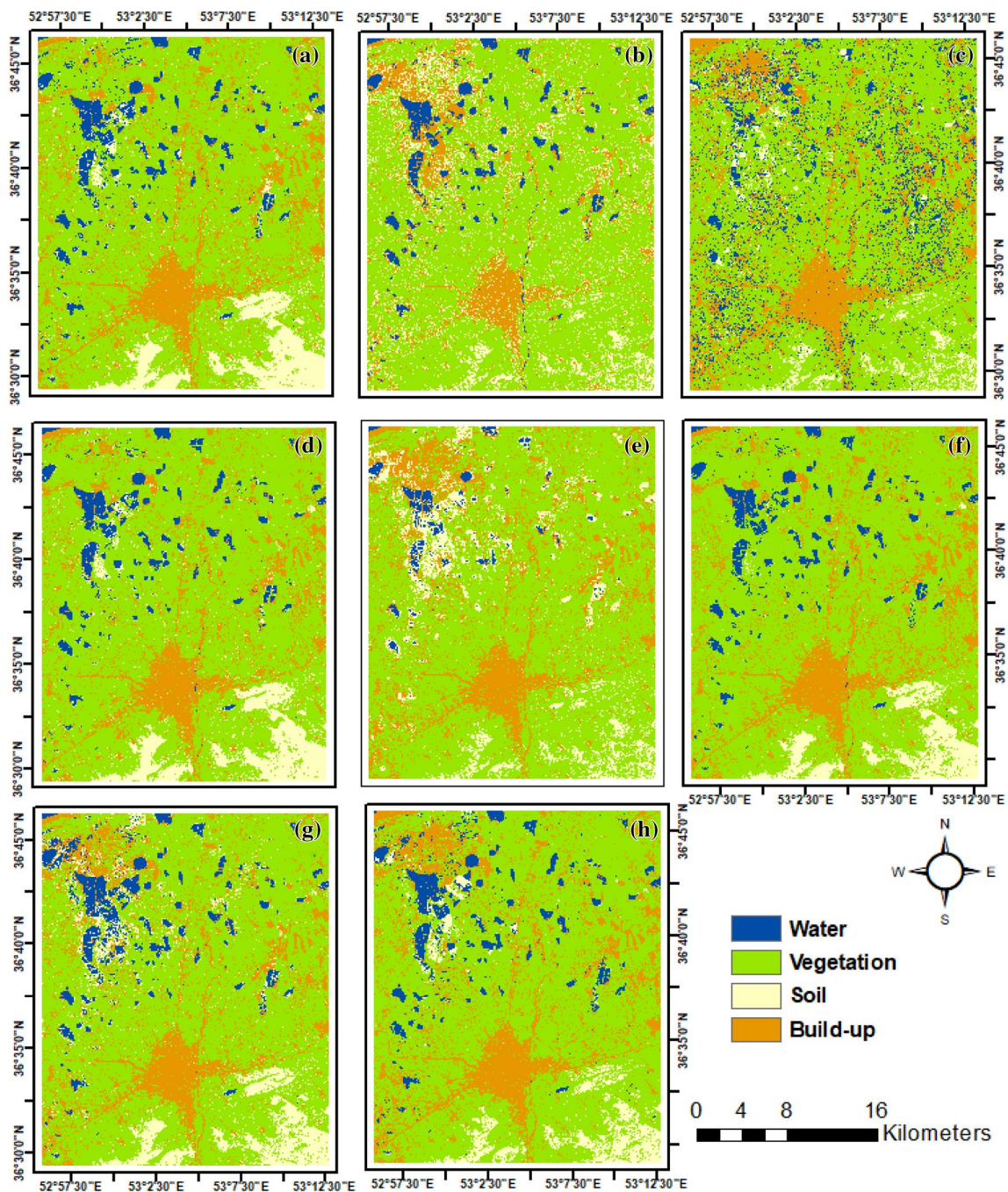
Fig. 4 Results of classification algorithms including **a** Multilayer Perceptron **b** Decision Table **c** DTNB **d** Lazy IBK **e** J48 **f** NN ge **g** Random Forest **h** Simple Logistic

52,776 cells as soil regions, 637,531 as vegetation regions and 39,587 cells as water regions. Random Forest classifier classified 148,874 cells as Build-up regions, 84,802 cells as soil regions, 595,609 as vegetation regions and 49,763 cells as water regions. 169,290 cells were classified as Build-up regions, 52,783 cells as soil regions, 609,559 as vegetation regions and 47,416 cells as water regions by Simple Logistic. Lazy IBK classifier classified 139,731 cells as Build-up regions,

74,392 cells as soil regions, 629,992 as vegetation regions and 38,933 cells as water regions.

6 Conclusions

Recently, due to the large-scale accessibility of Landsat imagery, the usage of LULC maps are highly increased. The LULC maps can be useful in order to retrieve fine-scale thematic data over zones confirming spatial analysis in various real-world bases including updating the road network, urban monitoring, etc. In the case of large area environmental analyzing, free and commercial Earth Observation (EO) satellites sensor information are used. Recently, image classification methods in large area environments are highly considered by researchers as a key factor because of various climate change impacts including increasing the temperature of the earth due to pollutant emissions like carbon dioxide (and also their influence on the land cover change). From the above discussion, it can be suggested that a fit-for-purpose algorithm should be suggested for a certain application such as vegetation extraction and flood modelling as well as man-made zone prediction. For the mentioned objective, WEKA and R programming language were used to implement and evaluate eight advanced machine learning algorithms including Random Forest, Decision Table, DTNB, J48, Lazy IBK, Multilayer Perceptron, NN ge, and Simple Logistic to classify Sari into four classes of water, vegetation, soil, and build-up regions. Results of this research can be used as a tool for monitoring forest regions in north of Iran where deforestation is a concerning issue. Iran's natural resources engineering office published an official report that Iran had 18 million hectares of forests in total, 3.4 m hectares of which were located in northern provinces of Gilan, Mazandaran and Golestan where the recent report indicates that forest areas have decreased to 14.2 million hectares and 1.8 m hectares in all three northern provinces. According to the presented results based on the training and test dataset, NN ge classifier which had the best performance in terms of OA, MAE, and RMSE is suggested as a fit-for-purpose algorithm to monitor deforestation in northern Iran within a time-series framework of LULCC mapping.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Bojinski S et al (2014) the concept of essential climate variables in support of climate research, applications, and policy. Bull Am Meteorol Soc 95:1431–1443
2. Shih H et al (2016) Determining the type and starting time of land cover and land use change in southern Ghana based on discrete analysis of dense landsat image time series. IEEE J Sel Top Appl Earth Obs Remote Sens 9:2064–2073
3. Bégué A et al (2018) Remote sensing and cropping practices: a review. Remote Sens 10:99
4. Foody GM (2002) Status of land cover classification accuracy assessment. Remote Sens Environ 80:185–201
5. Vitousek PM (1994) Beyond global warming: ecology and global change. Ecology 75:1861–1876
6. Skole DL (1994) Data on global land cover change: acquisition assessment and analysis. In: Turner I, Editor WB (eds) Changes in land use and land cover: a global perspective. Cambridge University Press, Cambridge, pp 437–471
7. Betts R et al (2007) Biogeophysical effects of land use on climate: model simulations of radiative forcing and large-scale temperature change. Agric For Meteorol 142(2–4):216–233
8. Grippa T et al (2018) Mapping urban land use at street block level using OpenStreetMap, remote sensing data, and spatial metrics. ISPRS Int J Geo-Inf 7(7):246
9. Costa H et al (2018) Land cover mapping from remotely sensed and auxiliary data for harmonized official statistics. ISPRS Int J Geo-Inf 7(4):157
10. Otukei JR, Blaschke T (2010) Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. Int J Appl Earth Obs Geoinf 12:S27–S31
11. Duro DC, Franklin SE, Dubé MG (2012) A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. Remote Sens Environ 118:259–272
12. Mountrakis G, Im J, Ogole C (2011) Support vector machines in remote sensing: a review. ISPRS J Photogramm Remote Sens 66(3):247–259
13. Breiman L (1984) Classification and regression trees. Chapman & Hall/CRC, Boca Raton
14. Hua L et al (2017) A feature-based approach of decision tree classification to map time series urban land use and land cover with Landsat 5 TM and Landsat 8 OLI in a Coastal City, China. ISPRS Int J Geo-Inf 6(11):331
15. Breiman L (1996) Bagging predictors. Mach Learn 24(2):123–140
16. Rodriguez-Galiano VF et al (2012) An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS J Photogramm Remote Sens 67:93–104
17. Fisher PF, Unwin DJ (eds) (2005) Representing GIS. Wiley, Chichester
18. Vapnik V (1998) Statistical learning theory. Wiley, New York
19. Vapnik VN (1995) The nature of statistical learning theory. Springer, New York
20. Breiman L (2001) Random forests. Mach Learn 45(1):5–32
21. Shao Y, Lunetta RS (2012) Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points. ISPRS J Photogramm Remote Sens 70:78–87
22. Pelletier C et al (2017) Effect of training class label noise on classification performances for land cover mapping with satellite image time series. Remote Sens 9:173
23. Karantzalos K, Bliziotis D, Karmas A (2015) A scalab legeo spatial web service for near real-time, high-resolution land cover mapping. IEEE J Sel Top Appl Earth Obs Remote Sens 8:4665–4674
24. Foody GM, Mathur A (2006) The use of small training sets containing mixed pixels for accurate hard image classification: training on mixed spectral responses for classification by a SVM. Remote Sens Environ 103:179–189
25. NASA (2013) Landsat 7 science data user's handbook

26. Cetin M (2015) Determining the bioclimatic comfort in Kastamonu City. Enviro Monit Assess 187(10):640

27. Cetin M (2015) Evaluation of the sustainable tourism potential of a protected area for landscape planning: a case study of the ancient city of Pompeipolis in Kastamonu. Int J Sustain Dev World Ecol 22(6):490–495

28. Cetin M (2015) Using GIS analysis to assess urban green space in terms of accessibility: case study in Kutahya. Int J Sustain Dev World Ecol 22(5):420–424

29. Cetin M et al (2018) Mapping of bioclimatic comfort for potential planning using GIS in Aydin. Environ Dev Sustain 20(1):361–375

30. Cetin M, Sevik H (2016) Evaluating the recreation potential of Ilgaz Mountain National Park in Turkey. Environ Monit Assess 188(1):52

31. Cetin M, Sevik H (2016) Assessing potential areas of ecotourism through a case study in Ilgaz Mountain National Park. In: Butowski L (ed), pp 81–110

32. Cetin M et al (2018) Evaluation of the recreational potential of Kutahya urban forest. Fresenius Environ Bull 27(5):2629–2634

33. Cetin M et al (2018) A study on the determination of the natural park's sustainable tourism potential. Environ Monit Assess 190(3):167

34. Brown A et al (2005) A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation. J Hydrol 310:28–61

35. Yeom J, Han Y, Kim Y (2013) Separability analysis and classification of rice fields using KOMPSAT-2 high resolution satellite imagery. Res J Chem Environ 17:136–144

36. Karakizi C et al (2018) Detailed land cover mapping from multitemporal landsat-8 data of different cloud cover. Remote Sens 10(8):1214

37. Karakizi C, Vakalopoulou M, Karantzalos K (2017) Annual crop-type classification from multitemporal landsat-8 and sentinel-2 data based on deep-learning. In: Proceedings of the 37th international symposium on remote sensing of environment (ISRSE-37), 2017, Tshwane, South Africa

38. Breiman L (2001) Random forests. Mach Learn 45:5–32

39. Catani F et al (2013) Landslide susceptibility estimation by random forests technique: sensitivity and scaling issues. Nat Hazards Earth Syst Sci 13:2815–2831

40. Kohavi R (1995) The power of decision tables. In: European conference on machine learning. Springer, Berlin

41. Hall MA, Frank E (2008) Combining naive bayes and decision tables. In: FLAIRS conference 2008

42. Quinlan JR (1993) The Morgan Kaufmann series in machine learning, San Mateo

43. Aha DW, Kibler D, Albert MK (1991) Instance-based learning algorithms. Mach Learn 6(1):37–66

44. Govindaraju RS, Rao AR (2013) Artificial neural networks in hydrology, vol 36. Springer, Berlin

45. Sylvain R (2002) Nearest neighbor with generalization. University of Canterbury, Christchurch

46. Landwehr N, Hall M, Frank E (2005) Logistic model trees. Mach Learn 59(1–2):161–205

47. Chai T, Draxler RR (2014) Root mean square error (RMSE) or mean absolute error (MAE)? Geosci Model Dev Discuss 7:1525–1534