

Performance Evaluation Report for LCLU Based on SVM, ANN and MLE Supervised Classification Algorithms Using Machine Learning

G. Nagalakshmi¹, T. Sarath²

Assistant Professor, Dept. of CS, Rashtriya Sanskrit Vidyapeetha, Tirupati. agnl.lakshmi@gmail.com Assistant Professor, Dept.of CSE, SISTK, puttur. sarath.pallapalli1@gmail.com

Article Info Volume 83 Page Number: 5798 - 5802 Publication Issue: May - June 2020

Article History Article Received: 19 November 2019 Revised: 27 January 2020 Accepted: 24 February 2020 Publication: 17 May 2020

Abstract

The Land Cover/ Land Use (LCLU) are the words used by the researchers who aware of classifying the land of a particular instant. Mainly this LCLU is used interchangeably, their actual meaning is pretty simple land use denotes the purpose of the land servers, for illusion mining, agriculture, settlement etc. Land cover denotes to the ground which covers the surface, vegetation based on whether, bare soil, water etc. These LCLU classification is done by exhausting remote sensing and GIS (Geographical Information System). The remote sensing images like multispectral, hyperspectral, LANDSAT, Sentinel, world view etc., GIS procedures can be used on vector polygon layers, raster images or mixture of both characterizes the land classification or detecting the LCLU. The remote sensing images are taken in the form of the daily basis. This study presents how LCLU classification can be done with the use of machine learning algorithms and analyzation of results based on various machine learning algorithms.

Keywords: Land Cover/Land use, Remote Sensing Images, GIS, Machine Learning, SVM, ANN, MLE.

I. INTRODUCTION

The land cover/land use inspection was first done by L.D. stamp at Britain in 1930. The classification of land is done to identify the problems and develop the purposes on aerial photographs and is applied by kellong in 1940 [11]. However, the LULC information and classification with remote sensing data was formulated in USA by 1971. The committee composed and implemented at U.S In survey of geological. The land cover/land use classification [12] system is efficient to employ the altitude and orbit of remote sensing to encounter the criteria.

- 1. 85 percent of remote sensing data is categorized to find the accuracy of the LULC.
- 2. It has numerous classes that has equal accuracy.
- 3. The classification is done for different sensing image which has different time.
- 4. The wide areas classification is applicable.

- 5. The vegetation and different types of land cover is categorised for active.
- 6. Aggregation of categories must be conceivable.
- 7. The land is recognised for different uses.

Aland cover/land use cataloguing system devised from USGS to find the remote sensing [13]information. The major classes that surveys: municipal or residential land, water, urban or builtup land agricultural land, barren land, forest land, wetland, range land.

In May 2006 national remote sensing centre devised LULC system is classified with remote sensing data of INDIA. The primary objective was mapping process with classification scheme and procedural steps for interpretation and mapping so as to maintain standard operational procedures. The classes are as follows wastelands, built up, water bodies, agricultural land, wetlands, forest etc.

Classification of LCLU is a criterion job for labelled indication of surface object contains the

pixel image in remote sensing. The complex structure of land use can cover different elements in land. The national mapping agencies will maintain the information about LULC and it stores in databases of geospatial. The main aim of LCLU is to describe the represented physical land type or how a land area is used. To full fill the above mentioned data firstly satellite images must be transformed into structured semantics. By seeing satellite image, we can observe different size and shape. Some may be readily identifiable while other may be not depending on our individual perceptions and experience. When we identify what we see on image and communicate the information to others, we are practicing interpretation.

A wide range of studies has been done on traditional parametric classifications [14] for remote sensing images and to produce high accuracy for the difficult data with a high dimensional feature space we will apply the machine learning algorithms.

II. MACHINE LEARNING ALGORITHMS

Machine learning algorithms are the prosperous development methods within the past few years. Some data demanding nominal and systematic fields alike search engines, robotics and speech recognition, object detection and soon. Over the few eras, studies have confirmed that a solitary data cause for land cover classification here, exit identification problem in traditional classification owing to lack of spectral data. Consequently, to increase the correctness of classification in dissimilar surroundings many readings have integrated in machine learning algorithm. The main three types of learning algorithm are as follows

Leaning of supervised

The procedure covers the goal/result/dependent variable that predicted in the form of predictors which is in sets (self- governing variables) via these set of variables, we make a function that map efforts to preferred productions. The guidance route stay still the model attains a needed level of accurateness on the exercise data. Illustration of supervised is decision tree, logistic regression, learning regression, random forest etc.

Learning of Unsupervised

In this procedure we don't have any goal/result variable to calculate/evaluate. It is worn for clustering populace in dissimilar group. This is commonly used for segmenting patrons in dissimilar collections for exact inversion. Illustration of learning of unsupervised is k means.

Learning of Reinforcement method

By means of this procedure, the machined proficient to sort explicit verdicts. It facility is that; the machine is uncovered to allocation anywhere it trains itself recurrently via testing slip. This machine learning from earlier experiences is try to imprison the finest feasible information to make exact commerce decision. For Illustration reinforcement learning, markov decision process.

III. STUDYAREA

To classify the image of satellite training set is trained and applied supervised classifier. These classifiers of supervised is parametric which needs preceding information about the pixel scattering like maximum likelihood estimator (MLE), Support vector machine (SVM) and Binary Decision Tree (BST) or non- parametric like Multi- layer perception, K-Nearest Neighbour and ANN.

The technique to classify the unsupervised utilizes clustering tools to collect the satellite image pixels into unlabeled clusters. Future predictor produces well classified image satellite by allocate sex pressive tags to the clusters [7]. Maximum classification of image satellite is by K-Means [2] which is unsupervised.

We will do comparison of numerous nonparametric and parametric classifiers. For land cover classification MLE is the most preferred because it uses highly resolution image. It is useful on the low resolution image like LANDSAT satellite image in another way non- parametric classifier Artificial Neural Network (ANN) is a learning procedure to assistant the meaning labels to image pixels. ANN algorithm can easily combine the supplementary data in the analysis process to improve the correctness of the image satellite [1]. It includes many binary rules to describe the expressive classes are related to specific pixel using Binary Decision Tree (BST).SVM is



Constantly identical in low resolution image because it uses different kernels so the performance will vary with the collection of kernel functions and its constraints [3].

K-Means is a widely held statistics and data mining method. Based on the Euclidean observations of n vales are partitioned into K clusters. These techniques have the advantage of modest process and fast execution. A drawback through this process is analysed to identify priori quantity of classes. Here we discuss different types of clustering algorithms that have mutual form of hierarchic clustering and accomplish with two deviations of fuzzy clustering.

In the supervised algorithm input data are trained using labelled, but in unsupervised data procedures are used against info which is not characterized. Complexity of learning supervised method is simpler method while other is computationally complex. Accuracy in supervised learning method is high but in unsupervised it is less accurate and both methods are trustworthy methods.

For comparison and to find which algorithm will give the average accuracy. For that taken image of SPOT 5 is used for analysis of experiment which is acquired from Suparco on 2013July. The image contains of infra-red, blue, red and green colours this image contains high resolution than Landsat image. Here datasets are distributed into many modules like tobacco, settled area, water, sparse vegetation, roads, barren land and sugar cane. After gathering datasets, training data is selected and ROI is created. Based on training set the data is trained and performed to evaluate follows as Figure 2 the User Accuracy (UA), Over All Accuracy (OA), Kappa Statistics (KS) and Producer Accuracy (PA). Based on ranges metrics are selected in remote sensing. OA affords the possibility of unidentified pixel to classify. UA determines the possibility of a pixel that planned in the specified class which it fits. PA defines the possibility of identifying the correct pixels. KS is percentage of the real accord by the allusion info beside the prospect union. Here collected the training data of all classes has the pixels of 36849. For classification each classifier is trained and calculated the four parameters. The classifier selection is built on their presentation in the land usage scrutiny and common classification chores.

Published by: The Mattingley Publishing Co., Inc.

Table 2 shows the PA and UA comparison for each ROI and for apiece classifier ANN, SVM, MLE is shown in Figure1. In that water class is reported as lowest UA based on SVM method with 53.8% which is highlighted as bold as mentioned in Table 2. Same as UA classes of water in ANN and MLE are stated 49.35 and 58.94 respectively. The similar situation was observed in Sparse Vegetation classification with concerning the parameters PA. In SVM [15] Sparse Vegetation has 71.74 of PA value whereas MLE has 81.18of PA via the same. For categorization Settled Areas of similar substitution be originating in SVM has 93.2 of PA, ANN [9] has92.3 of PA and MLE has94.9 of PA. The ROI are mentioned in the above are noted and recognised the good performance is highlighted in bold Sparse Vegetation, Settled Areas and Water and found MLE is coincided as good. Associated to the MLE, ANN only displayed the maximum accuracy for ROI in Tobacco, it has 94.83 of PA is presented in bold while, the MLE has91.65 of UA and SVM has 92.89of UA. Concerning the other ROI, UA has the highest for the Roads, Sugar Cane and Barren Land of SVM method as mentioned in the Table 2.

To estimate the whole presentation of classifiers SVM,MLE and ANN we mentioned the kappa statistics and overall accuracy. Table 3 express the KS and OA with the parallel classifier. As mentioned in table 3, we have noted that 84.8% of OA and0.8 of KS, the SVM [16] has outstripped all the classifiers. MLE [8] accomplishes well in ANN taking 83.2% of OA and 0.780f kappa statistics. Table 3 demonstrations that ANN [10] has the least OA of 80.4% and0.78 of KS.

Table 1. Details of each ROI Training set

Classifier	KS	OA
SVM	0.8	84.8
ANN	0.78	80.4
MLE	0.78	83.2

Table 2. Producer Accuracy(PA), User Accuracy(UA) evaluations of Performance

SVM		MLE		ANN		
	UA	PA	UA	PA	UA	PA
Road	81.74	70.95	73.99	72.75	73.44	59.92
Tobacco	92.89	85.05	91.65	84.39	94.83	78.33 5800



Barren	95.26	94.86	92.36	97.69	93.49	95.87
Land						
Tobacco	92.89	85.05	91.65	84.39	94.83	78.33
Sugar Cane	54.71	68.76	53.27	56.63	49.75	60.87
Sparse	79.11	71.74	81.18	68.1	61.26	75.5
Vegetation						
Settled	93.53	93.2	94.49	88.19	92.31	83.77
Areas						

By the above classification we analysed the land use based on the different methods of the supervised learning in machine learning algorithm using SPOT-5 image, supervised learning algorithms like ML, SVM, ANN We used 7 classes i.e. Roads, Tobacco, water, Settled Areas, Barren Land, Sparse Vegetation, and Sugar Cane. From the general investigation, it is determined that the SVM is fine stright to the land cover exploration. And also analysed it [4, 5, 6].



Fig 1. Classes detection of a land using SVM, ANN, MLE

IV. CONCLUSION

This article gives a summary on how the LCLU classification is done via machine learning algorithms and reviews done by various researchers. Here analyzed the different types of ML algorithms like learning of supervised method, learning of unsupervised method and learning of Reinforcement method. By these learning methods have different evaluations using SVM, MLE, ANN, DT etc., and compared the learning methods by taking an example satellite image, to that image training dataset is created using ROI and class are generated i.e. Roads, Vegetation, Water, Sugar Cane, Settled Areas, Tobacco and Barren land and evaluated data by using kappa statistics (KS), PA and UA. By analysing we conclude that SVM is best suited for the LULC by observing accuracy. So, finally we can also conclude that supervised classification will give high accuracy whereas unsupervised classification will give the less accuracy.



Fig 2. Work flow of the classification

 Table 3. Overall accuracy(OA) and Kappa Statistics(KS)Comparison

Classes	Description
Roads	Highways, other Routes used for Land
	Communication, Small Roads
Settled	Buildings, Restaurants, Hospitals,
Areas	Industries, Schools Shops, and Homes
Tobacco	Tobacco crop cultivated land.
Sparse	Contains of sparse vegetation like
Vegetation	Persimmon, Peaches, trees found on road
	side. Apples, leeches
Sugar	Sugar Cane crop cultivated land
Cane	
Barren	Land that cannot cultivate and not used
Land	for cropping.
Water	Streams, Canals, Reservoirs, Drains,
	Ponds, Channels, Rivers and Lakes.

REFERENCE

- [1] Kavzoglu, T., Reis, S. (2008). Performance analysis of maximum likelihood and artificial neural network classifiers for training sets with mixed pixels. GI Science and Remote Sensing, 45, 330–342.
- [2] Ahmed, R., Mourad, Z., Ahmed, B, H. and Mohamed, B.2009. "An Optimal Unsupervised Satellite image Segmentation Approach Based on



Pearson System and k-Means Clustering Algorithm Initialization", International Science Index, Vol. 3, No. 11, pp. 948-955.

- [3] Kavzoglu, T., Colkesen, I. (2009). A kernel functions analysis for support vector machines for land cover classification. International Journal of Applied Earth Observation and Geo information, 11, 352–359.
- [4] G.P. Petropoulos, K. Arvanitis, and N. Sigrimis, "Hyperion hyperspectral imagery analysis combined with machine learning classifiers for land use/cover mapping", presented at Expert Syst. Appl., 2012, pp.3800-3809.
- [5] Foody and Marther, 2004. A relative evaluation of multiclass image classification by support vector machine. IEEE Transactions on Geosciences and Remote Sensing.Vol 42.
- [6] Haung et al, Use of dark object concept and support vector machines to automate forest cover change analysis. Remote Sensing of Environment (2008). v112. 970-985.
- [7] Al-Ahmadi, F, S. and Hames, A, S. 2009. "Comparison of Four Classification Methods to Extract Land Use andLand Cover from Raw Satellite Images for Some Remote Arid Areas, Kingdom of Saudi Arabia", Journal of KingAbdul-Aziz University-Earth Sciences, Vol. 20, No.1, pp:167-191.
- [8] Scott, A.J. and Symons, M.J. (1971). Clustering methods based on likelihood ratio criteria, Biometrics 27(2), 387–397
- [9] Haykin, 1994.Neural networks. Macmillan College Publishing Company, New York, USA
- [10] J.F.Mas, J.J. Flores, The application of artificial neural networks to the analysis of remotely sensed data, International Journal of Remote sensing, Vol 29,No 3, PG 617-663.Feb 2008.
- [11] Gautam, A.P.; Webb, E.L.; Shivakoti, G.P.; Zoebisch, M.A. Land use dynamics and landscapechange pattern in a mountain watershed in Nepal. Agric. Ecosyst. Environ. 2003, 99, 83-96.
- [12] Peng, J.; Wang, Y.L.; Wu, J.S.; Yue, J.; Zhang, Y.; Li, W.F. Ecological effects associated with land use change in China's southwest agriculture landscape. *Int. J. Sustainable Dev. World Ecol.* 2006,13, 315-325.
- [13] Anderson, J.R., Hardy, E.E., Roach, J.T., Witmer, R. E. (1976). A Land Use and Land Cover Classification System for Use with Remote Sensor Data. United States Government Printing Office. Washington D.C.
- [14] Subhash, T., Akhilesh, S. and Seema, S.2012. "Comparison of Different Image Classification Techniques for Land Use Land

Cover Classification: An Application in Jabalpur District of Central India", International Journal of Remote Sensing and GIS, Vol. 1,Issue 1, pp. 26-31.

- [15] Haung et al, Use of dark object concept and support vector machines to automate forest cover change analysis. Remote Sensing of Environment (2008). v112. 970-985.
- [16] Yang, X. (2011) Parameterizing Support Vector Machines for Land Cover Classification, Photogrammetric Engineering and Remote Sensing, 77, 1, pp. 27-37.

Published by: The Mattingley Publishing Co., Inc.