A comparative analysis of machine learning algorithms for land use and land cover classification using google earth engine platform

Abhijit Patil[1]* and Sachin Panhalkar[1,2]
[1]Department of Geography, Shivaji University, Kolhapur, India.
[2]Center for Climate Change and Sustainability Studies, Shivaji University Kolhapur, India
*Email: abhijitpatil8893@gmail.com

(Received: 4 June 2023; in final form 17 October 2023)

DOI: *https://doi.org/10.58825/jog.2023.17.2.96*

Abstract: This study evaluates different machine learning algorithms for land use and land cover classification using Sentinel-2 Level-1C data with 10-meter spatial resolution. The algorithms include Random Forest (RF), Classification and Regression Trees (CART), Support Vector Machines (SVM), Naive Bayes (NB), and Gradient Boosting (GTB). The classification was performed on the Google Earth Engine (GEE) platform. Results highlight variations in land cover classification among algorithms, with RF and CART identifying cropland as dominant, SVM indicating fallow land presence, NB revealing significant forest cover, and GTB emphasizing cropland importance. Accuracy assessment was performed to evaluate the performance of the algorithms, considering metrics such as producer accuracy, consumer accuracy, overall accuracy, and Kappa coefficient. SVM demonstrates the highest overall accuracy and agreement with reference data. The study contributes insights for land management and planning, and GEE proves valuable for LULC classification.

Keywords: LULC, GEE, Machine Learning, Kolhapur

1. Introduction

Land Use and Land Cover (LU/LC) classification is a pivotal aspect of understanding the dynamic changes occurring on Earth's surface. Over the years, these changes have been influenced by a multitude of factors, including urbanization, agricultural expansion, industrialization, and climate variations (Turner et al., 2015). Accurately mapping and monitoring these transformations are crucial for informed decision-making in fields such as urban planning, environmental management, and natural resource conservation (Foody et al., 2013). This paper aims to shed light on the evolution of LU/LC classification methodologies, highlighting the pivotal role of remote sensing technology and the recent advancements in machine learning algorithms.

The advent of remote sensing technology has revolutionized the field of LU/LC classification. It has enabled researchers to collect and analyze vast amounts of data, providing a comprehensive view of Earth's surface across various temporal and spatial scales (Xie et al., 2019). Remote sensing data, acquired through satellites such as Landsat, MODIS, and Sentinel, have become invaluable sources for extracting critical information about land cover and its changes over time (Roy et al., 2014; Wulder et al., 2016; Gómez et al., 2016). These datasets offer a rich and diverse source of information, facilitating detailed analysis and monitoring of LU/LC patterns.

Over the past few decades, researchers have successfully harnessed the power of remote sensing imagery to create accurate and up-to-date LU/LC maps, capitalizing on its wide availability, comprehensive coverage, and ease of use (Stromann et al., 2019). However, generating such maps for large regions has presented challenges, including the need for substantial data storage, processing capacity, and diverse analytical approaches (Xie et al., 2019). Addressing these challenges led to the emergence of platforms like Google Earth Engine (GEE).

Before the introduction of GEE, various platforms were employed for LU/LC mapping, each with its own set of limitations. For example, while some platforms offered extensive datasets, they often lacked the computational power required for efficient processing (Shelestov et al., 2017; Pimple et al., 2018). Others provided robust processing capabilities but had limited access to high-quality remote sensing data sources. These constraints hindered the ability to produce accurate and timely LU/LC maps.

GEE was introduced as a game-changer in the realm of LU/LC mapping. This cloud-based platform seamlessly integrates an extensive collection of Earth observation data from various sources with high-performance computing services, enabling efficient processing of satellite imagery (Gorelick et al., 2017; Sidhu et al., 2018; Kolli et al., 2020). GEE offers access to a wide array of freely available satellite imagery, including datasets from sources such as Landsat, Sentinel, and MODIS (Shelestov et al., 2017; Pimple et al., 2018). It provides user-friendly interfaces in JavaScript and Python, facilitating the development and implementation of customized algorithms for satellite imagery processing and LU/LC mapping.

In light of these considerations, this study aims to leverage the power of GEE and conduct a comparative analysis of various machine learning algorithms for LU/LC classification using Sentinel-2 Level-1C data with a 10-meter spatial resolution. The research focuses on elucidating how different algorithms capture LU/LC variations and contribute to the advancement of mapping techniques. It seeks to address the critical question of which machine learning algorithm achieves the highest

accuracy in LU/LC classification, offering valuable insights for land management, planning, and environmental monitoring.

In this context, it is essential to explore and compare the performance of different machine learning algorithms in LU/LC classification (Level 1), with a focus on the area of interest in the present study, the Kolhapur City Region. By assessing the advantages and limitations of these algorithms, this research seeks to provide valuable insights for land management, planning, and environmental monitoring.

Through this investigation, we intend to demonstrate not only the efficacy of machine learning algorithms but also the importance of platforms like GEE in harnessing the full potential of remote sensing technology for LU/LC classification in a dynamic world.

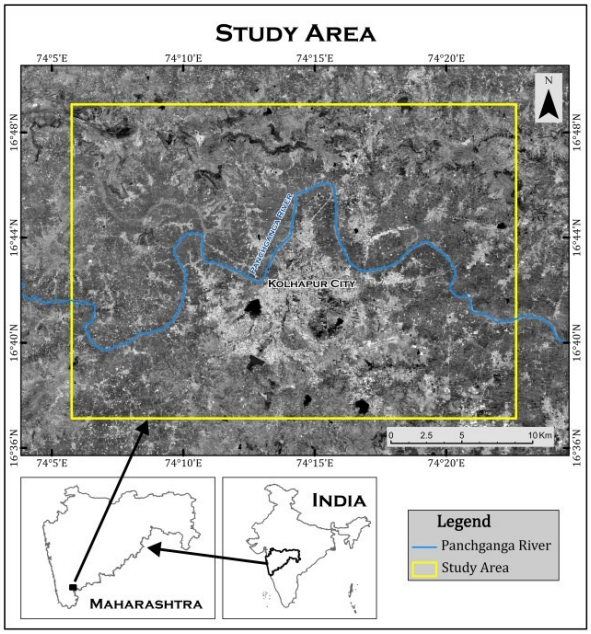

Figure 1. Location of study area

2. Study Area

The study area selected for this research comprises the surrounding region of Kolhapur city, situated in the south-western part of Maharashtra state, India. This region extends between 16° 37' 8" to 16° 49' 4" North latitude and 74° 5' 45" to 74° 22' 39" East longitude (Figure.1), covering an approximate area of 662 km².

Kolhapur city, renowned for its historical and cultural heritage, is strategically located on the southern bank of the Panchganga River within this study area. It serves as the central urban hub, experiencing rapid growth and urbanization. The choice of Kolhapur as the area of interest in this study is driven by its significance as a dynamically evolving urban center in the region.

In addition to the urban and agricultural areas, the study area includes significant natural features such as forests, water bodies, and barren land. These natural areas play a vital role in biodiversity conservation and maintaining ecological balance by providing habitat for a diverse range of plant and animal species.

The study area experiences distinct seasonal variations in rainfall, with the monsoon season typically extending from June to September. The annual average rainfall of this region is 1010mm. It is important to note that rainfall is a pivotal factor influencing land use and land cover dynamics in the region. It impacts agricultural practices, water availability, and natural ecosystems.

Regarding the physiography of the study area, it includes a diverse range of landscapes, from urban and agricultural plains to natural features such as forests and water bodies. The rapid growth of Kolhapur city and the surrounding region contributes to the changing physiography, making it an ideal location for studying land use and land cover dynamics.

3. Data and Methods

The systematic flow of the research methodology is depicted in Figure 2, illustrating the sequential steps undertaken in the study.

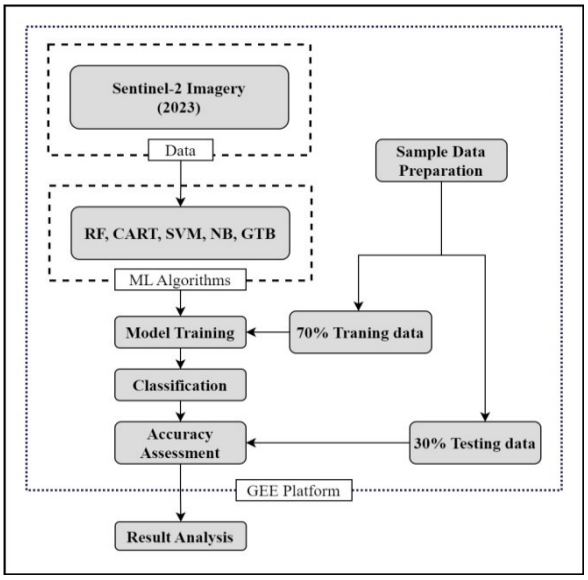

Figure 2. Research methodology

3.1 Data

The cloud-based GEE platform stores an extensive collection of Earth observation data (EOD) spanning over the past four decades. This vast dataset includes satellite images from platforms like Sentinel, Landsat, and MODIS, as well as other geospatial data such as climate and demographic information. Within the GEE platform, users have access to a wide range of EOD, including Sentinel data. This comprehensive data repository allows researchers to leverage a diverse array of satellite imagery and ancillary data to analyze and understand various aspects of the Earth's surface.

In the present study, Sentinel-2B (MSI) Level-1C satellite imagery from 05-03-2023 was utilized. To ensure data quality, only datasets with minimal cloud cover, constituting less than 10 percent of the total data, were

selected for analysis. For image classification purposes, three specific bands, namely Green (B3), Red (B4), and Near-Infrared (B8), from the Sentinel-2 dataset were employed. These bands have a spatial resolution of 10 meters, allowing for detailed and accurate classification of land use and land cover patterns within the study area.

3.2 Sample Selection

In supervised learning, training data plays a vital role, and a significant number of samples are typically required for most machine learning algorithms. However, obtaining accurate reference data from satellite imagery can be a challenging and complex task (Chi et al., 2008). The process of delineating and acquiring reference data involves identifying and labeling specific land cover classes within the satellite imagery, which can be a time-consuming and labor-intensive process.

In recent literature, the practice of utilizing separate training and testing datasets is widely adopted in the development of LULC classification models. Typically, a common ratio of 70% training data and 30% testing data is followed (Naceur et al., 2022; Huang et al., 2022; Wang et al., 2020). These reference sample points were generated based on high-resolution Google Earth images. For the current study, a total of six land cover classes were identified, and a total of 300 sample points were collected, with 50 sample points per class. Out of these, 35 sample points (Figure 3a) were designated for training the models, while the remaining 15 sample points (Figure 3b) were kept aside for the purpose of accuracy assessment. This practice ensures that the models are trained on a substantial amount of data, allowing them to learn the patterns and characteristics of the land cover classes. The testing dataset, comprising unseen data points, is used to evaluate the performance and generalization capability of the trained models.

3.3 Methods for Classification

The present study utilized the GEE platform to train classifiers for Sentinel-2 imagery by employing five different machine learning algorithms: Random Forest (RF), Classification and Regression Trees (CART), Support Vector Machine (SVM), Naive Bayes (NB), and Gradient Tree Boosting (GTB).

CART is a binary decision classification tree algorithm developed by Breiman (Breiman et al., 1984). It is a popular machine learning algorithm that enables straightforward decision-making in logical if-then scenarios. CART is a decision tree-based algorithm that recursively splits the data based on selected features to create a tree structure for classification or regression tasks. CART is advantageous for its interpretability and simplicity in visualizing and understanding the decision-making process. In the current study, the "*ee.Classifier.smileCart*" technique, GEE library, was employed to perform CART classification.

RF is an ensemble learning algorithm that combines multiple decision trees to make predictions. It is widely recognized for its ability to handle high-dimensional data and capture complex relationships between features and classes. The algorithm was first proposed by Leo Breiman (Breiman, 2001). One of the main advantages of Random Forest is its ability to reduce overfitting and improve generalization. In RF, the number of parameters and trees are two important factors that significantly impact the performance of the algorithm. These parameters are user-defined and can be adjusted based on the specific problem and dataset. The literature suggests that the optimal number of trees in a Random Forest typically ranges from 100 to 500 (Belgiu & Dragu, 2016). Having a larger number of trees can improve the model's accuracy, but there is a diminishing return as the number of trees increases. In the present study, 300 trees were selected as a compromise between computational efficiency and achieving a sufficiently accurate classification. In the current study, "*ee.Classifier. smileRandomForest*" technique as a GEE library, was employed to perform RF classification.

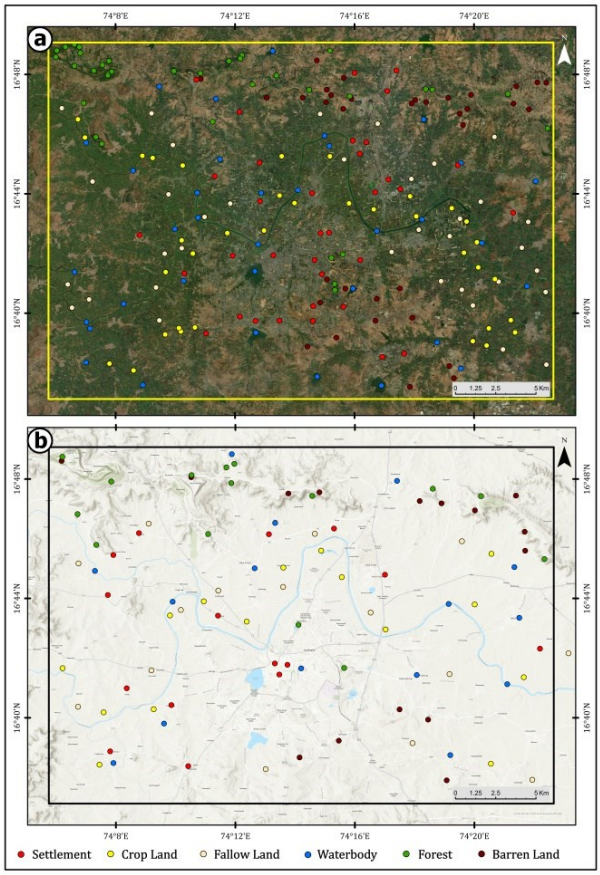

Figure 3. Training Samples (a) Testing Samples (b)

SVM is a machine learning technique that leverages the statistical learning theorem to identify an optimal higher-dimensional space by mapping the original input space. SVM finds an optimal hyperplane to separate different classes by maximizing the margin between them. This technique was initially introduced by Vapnik (Vapnik, 1995) and further developed by Cortes and Vapnik (Cortes & Vapnik, 1995) in the same year. The main parameters for selecting support vectors are the cost parameter C, Gamma, and kernel functions (Hsu et al., 2003). SVM offers several advantages for LULC classification. It can effectively handle non-linear relationships between input features and land cover classes by utilizing kernel

functions. In the current study, "*ee.Classifier.libsvm*" technique as a GEE library, was employed to perform SVM classification.

NB is a probabilistic classifier that assumes independence among features given the class and calculates the probability of each class given the input features using Bayes' theorem. Despite its simplicity, NB can perform well in various classification tasks, especially when the independence assumption holds reasonably well (John & Langley, 1995). NB is computationally efficient and can handle large datasets with high-dimensional feature spaces. It remains a popular choice for classification tasks, including LULC classification, due to its computational efficiency and effectiveness in certain scenarios. In the present study, "*ee.Classifier.smileNaiveBayes*" technique as a GEE library, was employed to perform NB classification.

GTB is a machine learning algorithm proposed by Jerome Friedman (Friedman, 1999). GTB is a boosting algorithm that combines multiple weak classifiers, usually decision trees, to create a strong ensemble model. GTB iteratively adds new trees to correct the errors made by previous trees, resulting in improved predictive accuracy. GTB is particularly effective in handling complex relationships and can capture non-linear interactions between features. In the present study, "*ee.Classifier.smileGradientTreeBoost*" technique as a GEE library, was employed to perform GTB classification.

3.4 Accuracy Assessment

Accuracy assessment is an essential step in evaluating the performance of LULC classification models. In this study, several metrics were used to assess the accuracy of the classification results, including producer accuracy, consumer accuracy, overall accuracy and Kappa coefficient. A total of 90 samples, with 15 samples for each class, were utilized for accuracy assessment purposes. These metrics allow for a comprehensive evaluation of the classification performance and provide insights into the reliability and effectiveness of the LULC classification models.

4. Result and Discussion

4.1 LULC Classification using GEE

This study investigated the performance of different machine learning techniques for land use and land cover (LULC) classification using Sentinel-2 Level-1C data with a spatial resolution of 10 meters. Figure 4 demonstrate how machine learning algorithms such as RF, CART, SVM, NB and GTB were used for the classification of LULC maps for 2023 on the GEE platform. The LULC classification focused on six major classes: settlement, cropland, fallow land, waterbodies, forest and barren land. The RF classification revealed that cropland occupied a significant land area of 266 square kilometers, accounting for 40% of the total area. This highlighted the agricultural dominance in the study area. However, compared to the other algorithms, RF had the second lowest land area for cropland. Waterbodies accounted for only 8 square kilometers (1% of the total area), indicating limited water resources in the study area. The CART classification showed an even larger land area dominated by cropland, covering 371 square kilometers (56% of the total area). This result further emphasized the agricultural dominance. Similar to the RF classification, waterbodies accounted for only 7 square kilometers (1% of the total area). The SVM classification identified a higher land area for the fallow land class, covering 120 square kilometers (18% of the total area). This indicated a significant presence of fallow land in the study area. Conversely, the settlement class had a relatively smaller land area of 99 square kilometers (15% of the total area). The NB classification produced a higher land area for the forest class, covering 219 square kilometers (33% of the total area). This indicated a significant presence of forests in the study area. However, the fallow land class had a relatively smaller land area of only 14 square kilometers (2% of the total area). It is important to note that the NB algorithm tends to classify fallow land into other classes, resulting in a lower land area for this category. The GTB classification revealed that cropland occupied a significant land area of 326 square kilometers (49% of the total area), making it the highest among all the algorithms for this class. Waterbodies accounted for 9 square kilometers (1% of the total area), representing the lowest land area among all the algorithms for this class.

The results of the LULC classification using different machine learning algorithms provided valuable insights into the land cover distribution in the study area (Figure 5). The RF and CART classifications consistently identified cropland as the dominant land cover class, highlighting the agricultural dominance in the region (Table 1). The SVM classification showcased a significant presence of fallow land, indicating land management practices in the study area. The NB classification revealed a substantial forest cover, while the GTB classification confirmed the importance of cropland in the landscape.

These variations in land cover classifications among the algorithms can be attributed to their inherent differences in modeling approaches and parameter settings. Each algorithm has its strengths and weaknesses in capturing specific land cover characteristics. Therefore, the choice of algorithm should be carefully considered based on the research objectives and the characteristics of the study area.

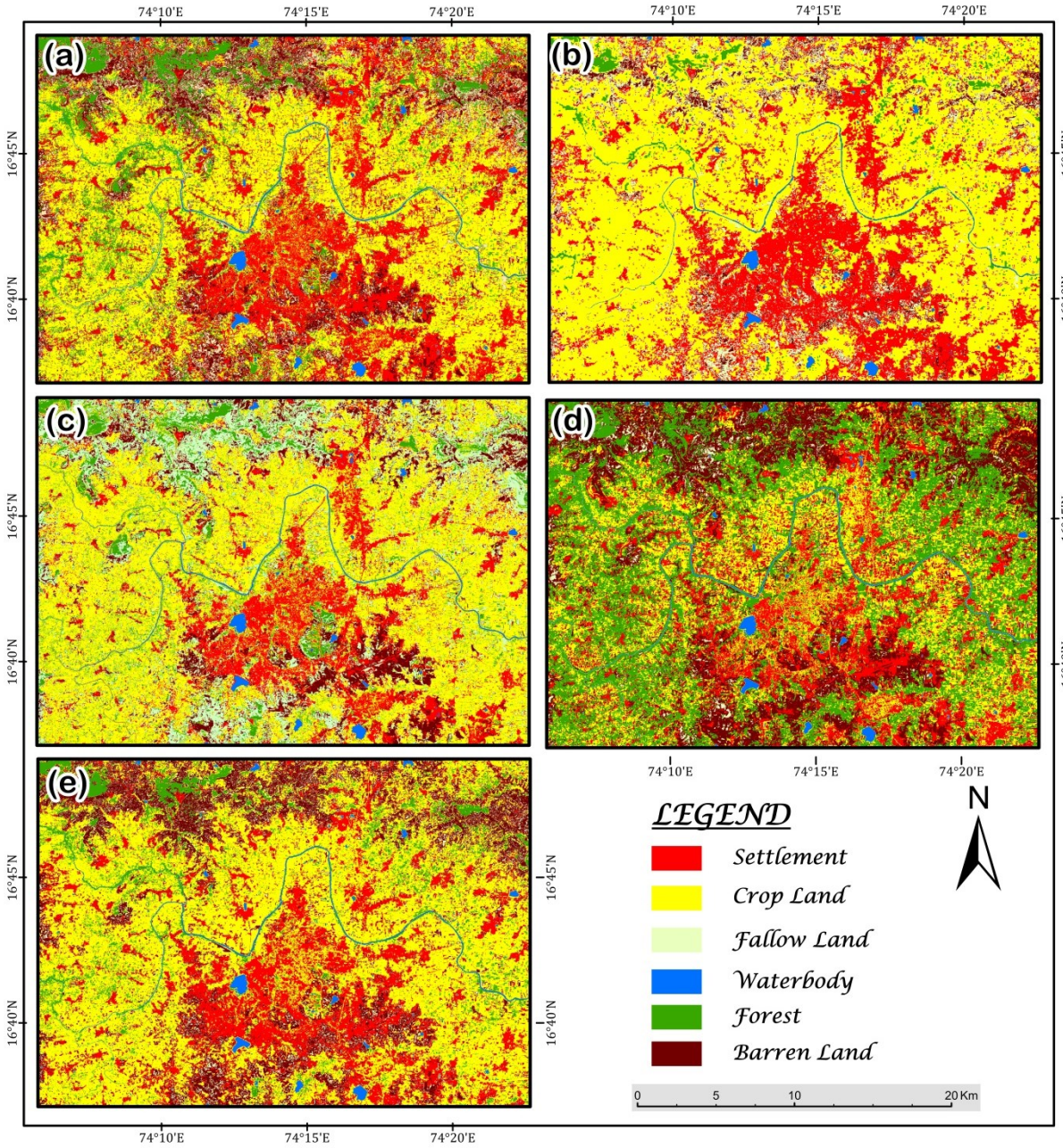

Figure 4. LULC classification (a) RF (b) CART (c) SVM (d) NB (e) GTB

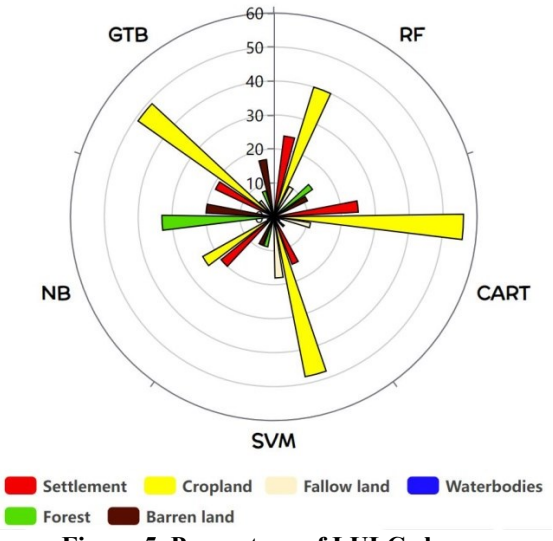

Figure 5. Percentage of LULC classes

4.2 Accuracy Assessment

The accuracy assessment was conducted to evaluate the performance of the machine learning algorithms in classifying land use and land cover (LULC) categories. The assessment involved calculating producer accuracy and consumer accuracy for each class, as well as overall accuracy and Kappa coefficient. Producer accuracy measures the correctness of classifying pixels belonging to a specific class. The results of the producer accuracy analysis for each algorithm are presented in Table 2. For the settlement class, RF, CART, SVM, and NB algorithms achieved high producer accuracy scores of 0.93, indicating accurate classification of settlement areas. GTB algorithm achieved a perfect producer accuracy score of 1. Cropland classification showed varied results among the algorithms. RF and SVM achieved high producer accuracy scores of 0.93, indicating accurate classification. CART had a relatively lower score of 0.8, while NB and GTB showed scores of 0.66 and 1 respectively.

Table.1: Percentage of LULC classes

Class/Algorithm	RF	CART	SVM	NB	GTB
Settlement	24.42	25.00	14.98	19.59	18.84
Cropland	40.16	56.10	47.96	24.43	49.27
Fallow land	9.84	11.01	18.18	2.06	5.93
Waterbody	1.14	1.05	1.13	1.18	1.32
Forest	13.81	2.40	8.84	33.07	7.59
Barren land	10.63	4.44	8.90	19.67	17.05
Total	**100**	**100**	**100**	**100**	**100**

Fallow land classification showed consistent high producer accuracy scores across most algorithms, ranging from 0.86 to 0.93. SVM achieved a perfect score of 0.93. Waterbodies were accurately classified by all algorithms, with producer accuracy scores ranging from 0.86 to 0.93. Forest classification results varied among the algorithms, with RF and SVM achieving high producer accuracy scores of 0.93 and 1 respectively. CART and NB showed lower scores of 0.73 and 0.53, while GTB achieved a perfect score of 1. Barren land classification had varied results, with RF achieving a producer accuracy score of 0.8, CART scoring 0.33, SVM scoring 0.93, NB scoring 0.86, and GTB scoring 0.6.

Consumer accuracy measures the correctness of classifying pixels that were labelled by reference data. The results of the consumer accuracy analysis for each algorithm are presented in Table 3. For the settlement class, RF, CART, and SVM algorithms achieved perfect consumer accuracy scores of 1, indicating accurate classification of settlement areas. NB and GTB algorithms achieved scores of 0.77 and 0.93 respectively. Cropland classification showed varied results among the algorithms. RF and SVM achieved relatively high consumer accuracy scores of 0.87 and 0.93 respectively. CART had a lower score of 0.63, while NB and GTB showed scores of 0.55 and 1 respectively. Fallow land classification showed varied results, with consumer accuracy scores ranging from 0.59 to 0.93. SVM achieved the highest score of 0.93. Waterbodies were accurately classified by all algorithms, with perfect consumer accuracy scores of 1. Forest classification results varied among the algorithms, with RF achieving a consumer accuracy score of 0.87, CART scoring 0.78, SVM scoring 0.93, NB scoring 0.8, and GTB scoring 0.88. Barren land classification had varied results, with RF achieving a consumer accuracy score of 0.92, CART scoring 0.71, SVM scoring 0.87, NB scoring 0.86, and GTB scoring 0.9.

Overall accuracy provides an assessment of the overall correctness of the classification results, while Kappa coefficient measures the agreement between the classification results and the reference data. Among the algorithms, SVM achieved the highest overall accuracy score of 0.94, indicating a high level of accuracy in classifying the LULC categories. RF, NB, and GTB algorithms achieved overall accuracy scores of 0.91, 0.8, and 0.9 respectively (Figure 6). CART algorithm had a lower overall accuracy score of 0.76. SVM also achieved the highest Kappa coefficient of 0.93, indicating a substantial agreement between the classification results and the reference data. RF, NB, and GTB algorithms achieved Kappa coefficients of 0.89, 0.86, and 0.88 respectively. CART algorithm had a lower Kappa coefficient of 0.72 (Figure 6).

These accuracy assessment results provide valuable insights into the performance of different machine learning algorithms for LULC classification. The SVM algorithm demonstrated the highest accuracy and agreement with the reference data, indicating its effectiveness in accurately classifying the LULC categories in the study area.

Table 2: Producer accuracy

Class/ Algorithm	RF	CART	SVM	NB	GTB
Settlement	0.93	1	0.93	0.93	1
Cropland	0.93	0.8	0.93	0.66	1
Fallow land	0.93	0.86	0.93	0.86	0.93
Waterbodies	0.93	0.86	0.93	0.93	0.86
Forest	0.93	0.73	1	0.53	1
Barren land	0.8	0.33	0.93	0.86	0.6

Table 3: Consumer accuracy

Class/ Algorithm	RF	CART	SVM	NB	GTB
Settlement	1	1	1	0.77	0.93
Cropland	0.87	0.63	0.93	0.55	1
Fallow land	0.82	0.59	0.93	0.86	0.73
Waterbodies	1	1	1	1	1
Forest	0.87	0.78	0.93	0.8	0.88
Barren land	0.92	0.71	0.87	0.86	0.9

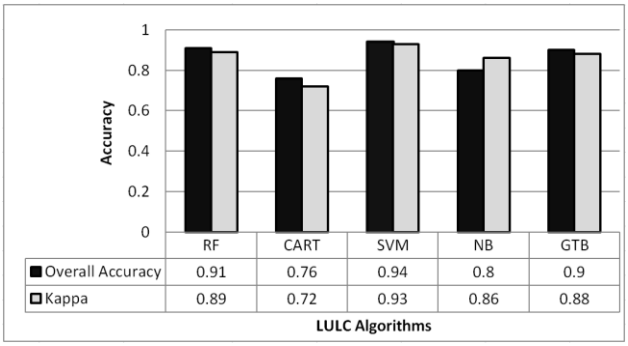

Figure 6. Overall accuracy and kappa coefficient

5. Conclusion

The primary objective of this study was to analyze the performance of various machine learning algorithms in the context of LU/LC classification using Sentinel-2 Level-1C data. Our findings shed light on the significant agricultural dominance within the study area, with cropland emerging as the dominant land cover class. Additionally, the study revealed the presence of fallow land, forests, and waterbodies, adding depth to our understanding of the landscape.

One of the key takeaways from this research is the pivotal role that the choice of algorithm plays in the classification outcomes. Our results underscore the importance of selecting the most suitable algorithm based on specific research objectives and the unique characteristics of the study area. It is evident that different algorithms excel in capturing distinct aspects of the landscape, which has implications for the accuracy and applicability of land cover classification results.

Furthermore, our accuracy assessment demonstrated the effectiveness of the SVM algorithm in accurately classifying the various land cover categories. This highlights the relevance of SVM in LU/LC classification tasks, but it should also encourage researchers to carefully consider algorithm selection for their own studies.
In conclusion, this research contributes valuable insights that hold significance for land management, urban planning, and environmental monitoring purposes. The dynamic interplay between machine learning algorithms and LU/LC classification results offers a fertile ground for further investigation and validation. Future research endeavors should aim to refine and expand upon these findings to enhance the accuracy and applicability of LU/LC classifications in diverse geographic regions.

Conflict of Interest
The authors confirm that there is no conflict of interest.

References

Belgiu M. and L.Dragu (2016). Random Forest in Remote Sensing: A Review of Applications and Future Directions. ISPRS Journal of Photogrammetry and Remote Sensing, 114, 24-31.

Breiman L. (2001). Random Forests. Machine Learning, 45, 5-32. https://doi.org/10.1023/A:1010933404324

Breiman L., J.H. Friedman, R.A. Olshen and C.J. Stone (1984). Classification and Regression Trees (1st ed.). London, UK: Routledge.

Chi M., R. Feng and L. Bruzzone (2008). Classification of Hyperspectral Remote-sensing Data with Primal SVM for Small-sized Training Dataset Problem. Advances in Space Research, 41(11), 1793-1799.

Cortes C. and V. Vapnik (1995). Support-Vector Networks. Machine Learning, 20(3), 273-297. https://doi.org/10.1007/bf00994018

Foody G.M., L. See, S. Fritz, M. Van der Velde, C. Perger, C. Schill, ... & F. Kraxner (2013). Assessing the accuracy of volunteered geographic information arising from multiple contributors to an internet based collaborative project. Transactions in GIS, 17(6), 847-860.

Friedman J.H. (1999). Greedy Function Approximation: A Gradient Boosting Machine. Annals of Statistics, 1189-1232.

Gómez C., J.C. White and M.A. Wulder (2016). Optical Remotely Sensed Time Series Data for Land Cover Classification: A Review. ISPRS Journal of Photogrammetry and Remote Sens, 116, 55-72.

Gorelick N., M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau and R. Moore (2017). Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone. Remote Sensing of Environment, 202, 18-27.

Hsu C.W., C.C. Chang and C.J. Lin (2003). A Practical Guide to Support Vector Classification. Technical Report, Department of Computer Science and Information Engineering, University of National Taiwan, Taipei, Taiwan.

Huang W., M. Ding, Z, Li, J. Zhuang, J. Yang, X. Li, L. Meng, H. Zhang and Y. Dong (2022). An efficient user-friendly integration tool for landslide susceptibility mapping based on support vector machines: SVM-LSM Toolbox. Remote Sensing, 14(14), 3408. https://doi.org/10.3390/rs14143408

John G.H. and P. Langley (1995). Estimating Continuous Distributions in Bayesian Classifiers. In Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence (pp. 338-345).

Kolli M.K., C. Opp, D. Karthe and M. Groll (2020). Mapping of Major Land-Use Changes in the Kolleru Lake Freshwater Ecosystem by Using Landsat Satellite Images in Google Earth Engine. Water, 12, 2493.

Naceur H. A., A.G. Hazem, I. Brahim, N. Mustapha, A. Hussein, A.A.D. Ahmed and A, Motrih (2022). Performance assessment of the landslide susceptibility modeling using the support vector machine, radial basis function network, and weight of evidence models in the

N'fis River basin, Morocco. Geoscience Letters, 9(1). https://doi.org/10.1186/s40562-022-00249-4

Uday P., S. Dario, S. Asamaporn, P. Sukan, L. Kumron, S. Henry , S. Jaturong , G. Valery and T. Sirintornthep (2018). Google Earth Engine Based Three Decadal Landsat Imagery Analysis for Mapping of Mangrove Forests and Its Surroundings in the Trat Province of Thailand. Journal of Computational Comm., 6, 247-264.

Roy D.P., M.A. Wulder, T.R. Loveland, C.E. Woodcock, R.G. Allen, M.C. Anderson, D. Helder, J.R. Irons, D.M. J ohnson, R. Kennedy, T.A. Scambos, C.B. Schaaf, J.R. Sc hott, Y. Sheng, E.F. Vermote, A.S. Belward, R. Bindscha dler, W.B. Cohen, F. Gao, J.D. Hipple and Z. Zhu (2014). Landsat-8: Science and Product Vision for Terrestrial Global Change Research. Remote Sensing of Environment, 145, 154-172.

Shelestov, A., L. Mykola, K. Nataliia, N. Alexei and S. Sergii (2017). Exploring Google Earth Engine Platform for Big Data Processing: Classification of Multi-Temporal Satellite Imagery for Crop Mapping. Frontiers in Earth Science, 5, 17.

Sidhu N., E. Pebesma E and G. Câmara (2018). Using Google Earth Engine to Detect Land Cover Change: Singapore as a Use Case. European Journal of Remote Sensing, 51, 486-500.

Stromann, O., A. Nascetti, O. Yousif and Y. Ban (2019). Dimensionality reduction and feature selection for object-based land cover classification based on Sentinel-1 and Sentinel-2 time series using Google Earth Engine. Remote Sensing, 12(1), 76.

Turner W., S. Spector, N. Gardiner, M. Fladeland, E. Sterling and M. Steininger (2015). Free and open-access satellite data are key to biodiversity conservation. Biological Conservation, 182, 173-176.

Vapnik V. (1995). The Nature of Statistical Learning Theory. https://doi.org/10.1007/978-1-4757-2440-0.

Wang G., L. Xinxiang, C. Chen, S. Himan and S. Ataollah (2020). Hybrid computational intelligence methods for landslide susceptibility mapping. Symmetry, 12(3), 325. https://doi.org/10.3390/sym12030325.

Wulder M.A., C.W. Joanne, R.L. Thomas, E.W. Curtis, S.W. Alan, B.C. Warren, A.F. Eugene, S. Jerad. G.M. Jeffrey and D.P. Roy (2016). The Global Landsat Archive: Status, Consolidation, and Direction. Remote Sensing of Environment, 185, 271-283.

Xie, S., L. Liu, X. Zhang, J. Yang, X. Chen, X. and Y. Gao (2019). Automatic Land-Cover Mapping Using Landsat Time-Series Data Based on Google Earth Engine. Remote Sensing, 11(23), 23.