Leveraging Machine Learning for Deforestation Detection and Sustainable Forest Management

Mohammad Shaazan Muzaffar[1], Utkarsh Singh[2], Zahabiya Hamied[3], Bharat Bhushan[4], Shivam Tiwari[5]

{m.shaazan@gmail.com[1], Trooftiger001@gmail.com[2], zahabiyahamied@gmail.com[3], bharat_bhushan1989@yahoo.com[4], shivam.tiwari.info@gmail.com[5]}

[1,2,3,4,5] Department of Computer Science and Engineering, Sharda University, Greater Noida, India

Abstract. Deforestation a global phenomenon which has been affecting the whole ecosystem, by not just effecting the biodiversity but the livelihoods of people relying on it as well. Recent studies have shown how this problem just tends on growing by each passing day, hence the need for predicting and detecting the drivers for this phenomenon becomes a crucial requirement for developing some responsive policies. The integration of Machine Learning (ML) techniques with the traditional methods of deforestation like monitoring and detection presents a promising solution for the sake of addressing the deforestation, the use of ML can help in preventing the various impacts of deforestation. The aim of this review is to provide a collective review of various ML algorithms and there uses across various dimensions of deforestation, highlighting both their potential and current limitations.

Keywords: artificial intelligence, deforestation, monitoring, biodiversity, supervised learning, unsupervised learning, sustainability.

1 Introduction

A significant environmental problem faced by various parts of the world is deforestation, as it disrupts ecosystems, favors climate change and loss of biodiversity. It is essential to monitor these changes by updating and accurately mapping how forest cover is structured (distribution) and reaction over time. In recent times remote-sensing technologies such as satellite observation, aerial drones and LiDAR measurements have emerged as critical instruments in detecting and monitoring forest degradation [1].

The growing use of Machine Learning (ML) and Deep Learning (DL) have contributed towards faster and more efficient analysis of such data. Techniques such as image segmentation helps to split images in the satellite images into disjoint regions, which can be used for easier identification of variation in vegetation instead clustered region and also an advantage on locating areas where forest logging is likely [2]. Trained on remote-sensing data, ML algorithms identify land features and detect anomalous changes, they can quantitate the size of cleared areas. Large areas can be scanned rapidly and accurately with these two capabilities [3].

Even with the overall advancements, a notable gap still exists. Present studies are primarily focused on large tropical forest tracts and very few such studies have been conducted in areas

IICTDS 2025, December 19-20, Chandigarh, India
Copyright © 2026 EAI
DOI 10.4108/eai.19-12-2025.2366216

undergoing rapid urbanization like Greater Noida. Vegetation loss in these regions also occurs more locally, with only small patches of plants being wiped out at a time, and is thus harder to model using the current models. There is still a lack of an integrated segmentation, classification and temporal evaluation framework specifically for such semiurban high-resolution datasets. Developing such an approach is essential for producing reliable, localized early-warning systems that can support better environmental planning and sustainable land-use decisions.

The major contributions of this paper are mentioned below:

• The paper presents a detailed overview of the various ML algorithms by classifying into 3 broad categories namely: Supervised, Unsupervised, and Reinforcement Learning.

• It categorizes the various aspects of deforestation, along with the ML implications for it.

• It highlights the various dimension of deforestation where ML can be applied and the recent advancements which can be made.

The paper has been organized into 5 sections which are as follows: Section 2 gives the detail about the different aspects of deforestation. Section 3 discusses about the various ML algorithms. Section 4 categorizes the different dimension of deforestation where ML can be applied and the recent advancements which can be made. Finally, Section 5 conclusion.

2 Aspects of Deforestation

Deforestation is among the direst concerns of today. Before we know the techniques for detection and monitoring of deforestation, we need to highlight the different aspects involved.

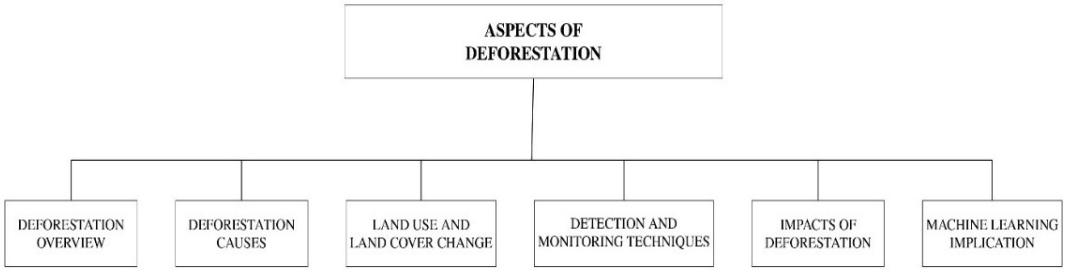


Fig. 1. Aspects of Deforestation.

2.1 Deforestation Overview

Deforestation is said to be the decline of forest areas across the globe. The main causes of this phenomenon are both the natural events and human activities, even though it is mostly pushed by humans via actions like urbanization, mining, agriculture farming. Forests play a vital role in the planet; Forests are providers that offers some of the most important components. Acting as a natural carbon sink, they eliminate the negative impacts of global warming by eradicating and storing CO2 in the atmosphere and are habitats for thousands of animals, birds, and plant species [4]. Different aspects of deforestration are shown in Figure 1.

2.2 Deforestation Causes

There are a number of natural events and different human activities which contribute to the rapid loss of green cover around the world resulting in deforestation. The main driver for the deforestation is the agriculture where 80% of the global forest has been converted to farmlands from forests to produce agricultural goods such as oil, rubber, soy, cocoa, coffee, and timber. The agricultural deforestation is driven because of increased demand of such different agricultural goods, other causes for the deforestation include mining, poor forest management, population growth, and illegal logging [5].

2.3 Land Use and Land Cover (LULC) Change

Forest Degradation and Deforestation are driven by changes in LULC. With each passing year, as the size of population grows, so does the need to shift forest land to accommodate agriculture, industrialization and urban expansion [6]. It does not only accelerate deforestation directly but also has secondary environmental effects, studies have shown. This change of use divides forest ecosystems, destroys the water cycle and increases natural disasters. Moreover, agricultural soil management practices worsen these issues by causing erosion, soil degradation and desertification, resulting in severe ecological and economic challenges, especially in fragile areas.

2.4. Detection and Monitoring Techniques

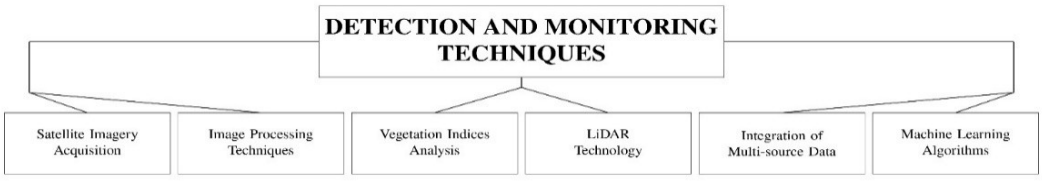


Fig. 2. Detection and Monitoring Techniques.

The use of remote sensing, Geographic Information System, and various other discoveries have resulted in evaluating environmental effects, guiding conservation initiatives, and monitoring deforestation trends as an effective tool. Differnt types of detection and monitoring are shown in Figure 2. The cycle of collecting data remotely via aerial platforms is called as remote sensing, visible light is used to collect data and produce images of earth's surface and various electromagnetic radiation like infrared, microwaves are used by these platforms. Some of these techniques have been mentioned below:

• Satellite Imagery Acquisition: Remote sensing primarily depends on satellite photography to monitor changes in forest cover overtime.

• Image Processing Techniques: Variety of image processing techniques like classification, change detection and spectrum analysis are used by GIS specialist to extract useful information from satellite photos.

• Vegetation Indices Analysis: The Normalized Difference Vegetation Index (NDVI) are often utilized in remote sensing. Areas experiencing degradation and deforestation can be pinpointed by GIS specialists by monitoring chains in NDVI values over time.

• LiDAR Technology: Highly accurate 3D dimensional data about the structure of forests and topography can be gathered by using light detection and raging LiDAR technology.

• Integration of Multi-source Data: Data is integrated by GIS specialist from various resources like satellite imaging, ground observations and LiDAR data to build complete deforestation monitoring systems. The accuracy and dependability of evaluations of deforestation is enhanced by this combination.

• ML Algorithms: By using advanced ML techniques like Support Vector Machines (SVM) and RF remotely sensed data is processed to automatically detect patterns of deforestation.

2.5. Impacts of Deforestation

On a local level the effect of deforestation includes the loss of valuable ecosystem services provided by forests. These habitats provide critical services to human including flood control, fisheries protection, erosion prevention, and pollination services. The loss of forest also reduces the presence of renewable resources. This can have global implication affecting the global climate and biodiversity.

• Mining: Extractive actions are required to clear forests. Due to mining operations both the Amazon and Congo basins in particular have seen large deforestation.

• Roads: Building routes facilitates deforestation by providing easy access to previously remote forest areas, often leading to increase in logging settlement and agriculture expansion.

• Urbanisation: To make way for new homes, commercial buildings and infrastructures, forest is being emptied, thus expanding urban areas.

• Agriculture: The leading cause of deforestation is expanding agriculture. Often to plant crops or graze livestock forests are emptied.

• Clearance for livestock: To create pastures for cattle, natural forests are cleared but these livestock do not just only need space to graze. To feed farmed animals, space is also needed to grow crops in animal agriculture.

• Timber: Widespread deforestation is caused by cutting down of trees. Illegal logging [7] is accounts for 15% to 30%, when it comes to global timber production.

• Fires: Fires, whether started naturally or by humans, contribute to deforestation. The destructive impact of uncontrolled fires in the Amazon Forest have drawn global attention for their disastrous impact.

2.6. Machine Learning Implication for Deforestation

To prepare future action plans, the most important inputs for forest management are deforestation identification and mapping. To estimate management approaches identifying and evaluating the changes in forests resources improve the forest management by providing framework to the stakeholders and policymakers.

In present time the forests are more exposed than in the past, as the population is increasing day by day, and for successful implementation of operations, projects relevant awareness is much needed. To develop an advanced forest monitoring system to minimize the cost ML technologies are really needed [8].

Among the various ML technologies, Remote Sensing, that provides important and precious information from a long range. For remote sensing private satellites such as Landsat and Sentinel 2 offer intensified spectral, spatial, and temporal resolution that provides open data. Sentinel-2 with its various multispectral bands that include the vegetation red edge bands which offer 5 days of temporal data for regular inspection. By using these technologies, machine learning enhances the accuracy and efficiency of forest type detection, deforestation pattern and monitoring of ecosystem health, hence contributing efficiently to sustainable forest management [9].

3 ML Algorithm Overview

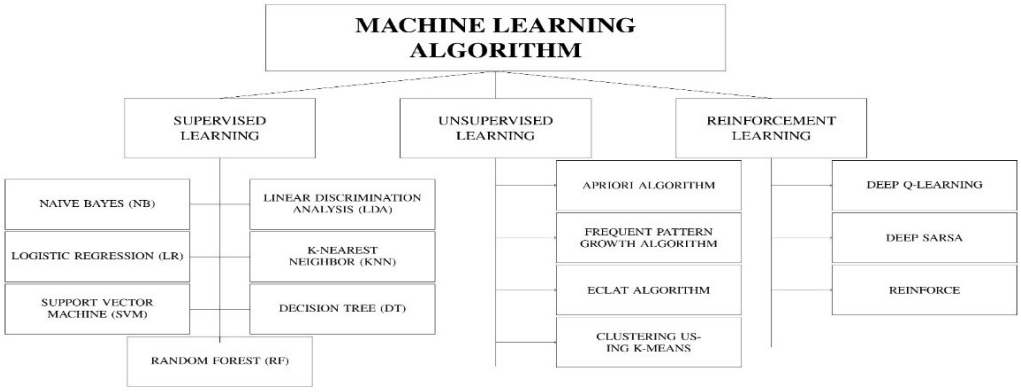


Fig. 3. ML Algorithms.

The algorithms of ML have been classified into different categories which have been briefly discussed below:

3.1 Supervised Learning

It uses sample pair of input-output to learn a mapping function that relates an output over an input. Different types of ML algorithms are shown in Figure 3. The supervised tasks which most common include “regression” and “classification” which separates the data [10].

Naive Bayes (NB): The algorithm in many real-world situations it can be used for both multi-class and binary categories and works well in situations like text or document classification, filtering spams etc., as shown in equation (1).

$$\hat{y} = arg \underset{c \in C}{max} P(C) \prod_{i=1}^{n} P(x_i|c) \quad (1)$$

Where, c represents class, n represents number of features, and xi represents feature i value.

Linear Discrimination Analysis (LDA): Also called as a generalization of Fisher's linear discriminant, the basic LDA model normally suits every class with Gaussian density separately, assuming that the same covariance matrix is shared among the all classes, as shown in (2).

$$\delta_k(x) = x^T \Sigma^{-1} \mu_k - \frac{1}{2} \mu_k^T \Sigma^{-1} \mu_k + l\eta \pi_k \quad (2)$$

Where, μk represents mean vector of class k, Σ represents shared covariance matrix, and πk represents prior probability of class k.

Logistic Regression (LR): A ML model that is used to get the answer for classification issues. It commonly uses a logistic function. It can overfit the complex datasets and work pretty-good when dataset is divided linearly. The presumption of linearity among the independent and dependent variables is regarded as a major weakness of LR, as shown in (3) and (4).

$$P(y = 1|x) = \frac{1}{1+e^{-(w^r x+b)}}, \quad (3)$$

$Where, x$ represents feature vector, ω represents weight vector , $and\ b\ represents\ bias.$

$$d(x) = \frac{w^r x+b}{||w||}\ (distance\ from\ decision\ boundary) \quad (4)$$

K-Nearest Neighbour (KNN): The model is a method for classification of new objects by measuring its distance from training samples in the feature space. It finds the k closest neighbours using distance metrics like Manhattan or Euclidean and assigns the majority class among these, employing weighted voting for neighbours. It requires no training process, is simple to implement, and performs well without any parameters, as shown in (5) and (6).

$$\hat{y} = mode\ \{y_{(1)}, y_{(2)}, \ldots, y_{(k)}\} \quad (5)$$

Where, k represents number of nearest neighbors, and yi represents class of the i-th nearest neighbor.

$$d(x, x_i) = \sqrt{\sum_j \left(x_j - x_{ij}\right)^2}\ \ (Euclidean) \quad (6)$$

Where, $d(x, x_i)$ represents Euclidean distance between query point x and i-th neighbor x_i, with j denoting feature index.

Support Vector Machine (SVM): The algorithm is derived on statistical learning theory and its main concept is to reduce errors in predictions. The decision-making of this model is focused on a small amount of data as a large number of useful patterns can be extracted, as shown in (7).

$$\min_{\omega,b} \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{m} max\left(0, 1 - y_i(\omega^T x_i + b)\right) \quad (7)$$

Where, m represents number of training samples, $yi \in \{-1,+1\}$ represents class label of sample i, and C represents regularization parameter.

Decision Tree (DT): It aims at identifying vital characteristics for a dataset by splitting the information into smaller groups that contain the best feature values. The model includes two nodes: parent node and the leaf nodes. The model makes predictions based on answers to conditions at each step as we move from root to leaf of the tree, as shown in (8) and (9).

$$IG(t,f) = H(t) - \sum_{v \in values(f)} \frac{|t_v|}{|t|} H(t_v) \tag{8}$$

Where, H(.) represents entropy of samples at node, t represents set of samples at node, and tv represents subset with feature f taking value v.

$$H(S) = -\sum_{i=1}^{n} p_i \, log_2(p_i) \tag{9}$$

Where, H(S) represents entropy of dataset S, n represents number of distinct classes, and p_i represents proportion of samples belonging to class i.

Random Forest (RF): RF algorithm is used for both regression and classification and operates through constructing DT. This model is intuitive, good at dealing with outliers, and performs better compared to a single DT, which is why it is commonly used in ML competitions. RF is a tree-ensemble method; predictions are gathered from many individuals, each trained separately. It can detect faults like degradation, open circuits, partial shading, and line-to-line faults in systems. Thus, N number of trees can be trained and utilized for predictions, as shown in (10).

$$\hat{y} = mode\ \{\hat{y}^{(1)}, \hat{y}^{(2)}, \ldots, \hat{y}^{(T)} \tag{10}$$

Where, T represents number of trees, and yj represents prediction by tree j.

3.2 Unsupervised Learning

Unsupervised learning is that kind of ML which is vital to generate human artificial intelligence systems. The principle of artificial neural network is also based on unsupervised learning [11].

Apriori Algorithm: The Apriori is employed to obtain repeated item sets from large trade-based databases such as in supermarkets to identify items often purchased together. It finds applications in student data, forest fire patterns, patient databases, as shown in (11).

$$Support\ (X) = \frac{|\{t|X \subseteq t\}|}{N} \tag{11}$$

Where, X represents itemset, t represents a transaction, and N represents total number of transactions.

Frequent Pattern Growth Algorithm: The Frequent Pattern Growth (FP-growth) Algorithm is a significant technique in association rule. By building a Frequent Pattern Tree (FP-tree) and applying a divide-and-conquer method, FP-growth avoids candidate generation, as shown in (12).

$$Support\ (X) = \frac{Number\ of\ transaction\ containing\ X}{Total\ number\ of\ transactions} \tag{12}$$

Where, X represents itemset, and N represents total number of transactions.

ECLAT Algorithm: The ECLAT algorithm is a data mining technique employed to obtain frequent itemset and evaluate transaction data optimally. It begins by listing transaction IDs for individual items, overlaps those lists to find common occurrences, and continues the process selectively with other items, as shown in (13).

$$Support\ (X \cup Y) = |TID(X) \cap TID(Y)| \ / \ N \tag{13}$$

Where, TID(X) represents set of transaction IDs containing itemset X.

Clustering Using K-Means: K-Means is a learning technique which results in clustering of unclassified data points based on similarities. Until the optimal positions are found. This non-hierarchical grouping method is faster than hierarchical clustering for large datasets with many variables, as shown in (14).

$$\min_{\{C_k\},\{\mu_k\}} \sum_{k=1}^{k} \sum_{x_i \varepsilon C_k} \|x_1 - \mu_k\|^2 \tag{14}$$

Where, K represents number of clusters, Ck represents cluster k, and μk represents centroid of cluster k.

3.3 Reinforcement Learning

Reinforcement learning is held to be an important part of ML. It enables an agent to develop effective behaviours by directly interacting with its environment and utilizing feedbacks in the form of responses. Positive feedbacks encourage beneficial decisions while it is important to carefully choose a model based on a specific environment for real world applications [10].

Deep Q-Learning (DQN): DQN networks are essentially reinforcement learning algorithms. By integrating convolutional and fully connected layers, DQN maps environment states to Q-values that guide action selection. Estimations are further improved by Dueling DQN. In double Dueling DQN the two approaches are combined to enhance performance. Together, these variants improve learning stability, action accuracy, and efficiency across complex, dynamic environments, as shown in (15).

$$L(\theta) = E_{(s,a,r,s')\sim D}\left[\left(r + \gamma \max_{a'} Q(s', a'; \theta^-) - Q(s, a; \theta)\right)^2\right] \tag{15}$$

Where, s represents state, a represents action, r represents reward, s' represents next state sample, γ represents discount factor, θ represents parameter of current, and θ− represents target networks.

Deep SARSA: Deep SARSA expands on the traditional SARSA algorithm. Actions are selected based on maximum Q-values. Q-values are updated using the next action taken. This approach leads to policy gradient methods. The policy in these methods is updated on the based on performances. In on-policy methods to evaluate and collect data same policy is used. On the contrary off-policy methods rely on a different behaviour policy to gather data, as shown in (16).

$$(\theta) = E_{(s,a,r,s',a')\sim D}[(r + \gamma Q(s', a'; \theta) - Q(s, a; \theta))^2] \quad (16)$$

Where, s represents state, a represents action, r represents reward, s' represents next state sample, γ represents discount factor, θ represents parameter of current, and a' represents action actually taken in next state.

REINFORCE: REINFORCE is an essential, sampling based, action-policy-derived algorithm used in reinforcement learning, particularly. It operates by collecting complete episodes. The key idea is to optimize the policy. It experiences high variance in gradient estimates. To address this, a reference value, which is one of the first policy gradient methods, as shown in (17).

$$\nabla_\theta J(\theta) = E_{T\sim\pi\theta}[\sum_{t=0}^{T} \nabla_\theta \ln \pi_\theta(a_t|s_t)\, G_t] \quad (17)$$

Where, θ represents policy parameters, T represents trajectory of states and action, and Gt represents cumulative reward from time t.

4 Dimension of Deforestation Where ML Can Be Applied

Deforestation is a phenomenon that is driven by environmental, social, and economic factors. Various applications of ML shown in Figure 4 that can be applied to several dimensions, which have been discussed below:

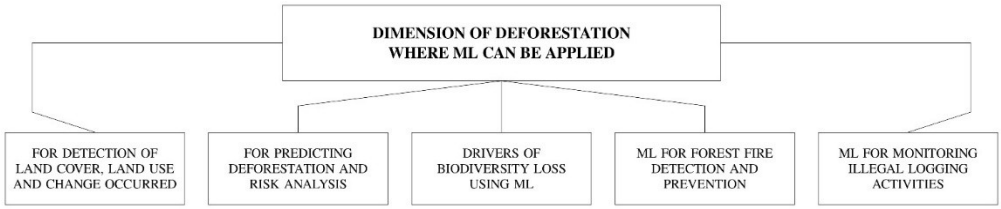


Fig. 4. Dimension of Deforestation where ML Can Be Applied.

4.1 ML Used for Detection of LULC and Change Occurred

The difficult part is to automate the LULC and detecting and changes occurred. In some related studies, Wang et al. [13] proposed a crowdsourcing collaborative framework and an automatic model for change detection. Similarly, Sofi et al. [14] presented a study which aims to analyse patterns of gait in a compiled dataset of a population in Morocco, for the purpose of classifying the gait models with the use of GAIT Rite system; with the help of machine learning (ML) and the refined gait parameters they detected gait patterns which are linked to various neurodegenerative disorders. In another work, Chintya et al. [15] developed a change detection

model which is based on deep learning for multitemporal images taken from UAV's (Unmanned Aerial Vehicles) for breakwater facility and SNUNet-CD algorithm, a mixture of Nested-U-net and Siamese Network, was used for change detection in the dissipating blocks of the facility, Serra et al. [16] highlighted that under unpredictable scenarios like occurring of double bands at same areas with various intensities or at different areas with same kind of intensities the physical model based ITCZ detection alone could misclassify they then designed some AI models trained on experiments to collect results of ITCZs detection using test automation., Li et al. [17] proposed Dyn-Watch which is topology-aware algorithm and domain knowledge based for the detection of anomaly with the help of sensors arranged on a dynamic grid.

Table 1. Summary Of ML Applications In Deforestation Dimensions.

Dimension of Deforestation	References	Year	Major Contributions	Limitations
Detection Of Changes Occurred	Wang et al. [13]	2024	Intelligent collaborative change detection enabled by Active learning crowdsourcing framework.	Needs many users
	Sofi et al. [14]	2024	Patterns of neurodegenerative disorders are linked by using ML gait analysis.	Not geospatial data
	Chintya et al. [15]	2022	By using SNUNET-CD UAV-based DL change detection for breakwater monitoring.	High computation
	Serra et al. [16]	2022	ITCZ detection improved beyond traditional approaches by AI models.	Less accurate in extremes
	Li et al. [17]	2022	Via sensor grids and topology aware modelling anomalies are detected by Dyn-Watch.	Grid-based only
	Shayesteh et al. [18]	2024	To manage feature drift and refine predictive model's frameworks are developed.	Frequent retraining
	Reddy et al. [19]	2024	To analyse the records of patient for prognostic factors applied DL is used.	Domain specific
Predicting and Risk Analysis	Rampal et al. [20]	2024	For better accuracy Multiple classifiers are integrated by creating hybrid soft voting model.	Complex model
	Chowdary et al. [21]	2023	By using ML models early cardiovascular prediction, outlier detection and preprocessing.	Sensitive to noise
	He et al. [22]	2020	By using diverse ML methods designed optimal customer churn prediction.	Not environmental

Drivers Of Biodiversity Loss Using ML	Molina-Rotger et al. [23]	2024	For insect monitoring Vio view, a solar power remote sensor system is developed.	Limited scope
	Lang et al. [24]	2024	Four species detection DL is used to propose autonomous bat echolocation system.	Species-specific
	Su et al. [25]	2024	Data is sensed for ecosystem analysis by employed ML on two-decade multisource monitoring.	Region limited
	Lopez-de-Ipina et al. [26]	2015	By using auditory based systems and ML environmental agents and interactions are studied.	Outdated methods
	wa Maina [27]	2015	For bird census and biodiversity monitoring signal processing and ML are jointly used.	Low robustness
Detection and Prevention	Thakur et al. [28]	2024	For personalized risk assessment ensemble learning model was proposed integrating multidata type.	Data-heavy
	Toyin et al. [29]	2024	To analyse real time network traffic ML based intrusion detection system was built.	False positives
	Giri et al. [30]	2024	By using high resolution image DL cervical cancer detection was developed.	Needs large data
	AlMasri et al. [31]	2022	To defend against DoS and scanning attacks ML pattern recognition and networking were used.	Misses new attacks
	Alam et al. [32]	2022	Achieved an accuracy of 99.62% by presenting an exceptional model for fraud detection.	Possible overfitting
Monitoring	Manimozhi et al. [33]	2025	To detect freshness and dishes automatically an AI driven nutrition monitoring was proposed.	Not outdoor-suitable
	Kalita et al. [34]	2024	For real time remote diabetes patient monitoring cloud-based ML system was developed.	Cloud dependent
	Chaudhry et al. [35]	2024	For real time Elderly Healthcare parameter assessment RF and SVM are used.	Sensor variability

Mishra and Singh [36]	2024	From variable health parameters ML model to predict fitness scores was built.	Low generalization
Ramu et al. [37]	2024	An air quality monitoring, Edge-AI, that's a decentralized with IoT's and Nvidia Jetson Is designed.	Costly hardware

4.2 ML for Predicting Deforestation and Risk Analysis

AI driven deforestation risk analysis can provide benefits to your business in several ways including: Recognising and reducing the risk of deforestation, keeping transparency across the supply chain, collaborating effectively with stakeholders. In some related discoveries, Shayesteh et al. [18] proposed a framework for managing feature drifts occurring in this environment. It detects the important features and uses the predictive model's performance to detect feature drift, and refines the predictive model and enhances the features along the drift depending on the degree of the feature drift. Similarly, Reddy et al. [19] identified the uses of a DL model which can be used to study large loads of patient records in order to identify factors and complex patterns facilitating to prognosis, Rampal et al. [20] mentioned a model which integrates various classifiers, such as SVM, RF, KNN, DT, and LR, with the use of soft voting assembly approach. The hybrid model being effective was evaluated with the help of accuracy metrics and classification reports. Also, Chowdary et al. [21] proposed an early prediction model for cardiovascular disease, where data pre-processing, outlier detection, predictive ML models like k-NN, NB, XGB, SVM were employed, He et al. [22] outlined an optimal model for the customer churn rate prediction while working on different ML methods. The dataset includes macro environmental features, customer behaviour features, and customer demographics features.

4.3 Drivers of Biodiversity Loss Using ML

There is a huge need of easy, accessible and interpretable methods that supports recursively and effectively analyse deforestation. In some related findings, Molina-Rotger et al. [23] introduced an automated electronic system made for monitoring the insect species population called Vio VIEW remote monitoring station. Equipped with several sensors, light, solar-powered supply system, that is installed in remote locations for its work. While, Lang et al. [24] proposed the Autonomous Bat Echolocation Monitoring System, equipped with auditory sensors, devices and DL algorithms to monitor bat movements and detecting and classifying its species as this system is for preservation and pest control in agriculture, Su et al. [25] highlighted the use the machine learning techniques and utilizes quantitative remote sensing, while using the data of Inner Mongolia from around two decades of multisource remote sensing and meteorological ranging from 2000 to 2017.Similarly, Lopez-de-Ipina et al. [26] focused on the study of environment agents as well their interactions using automatic systems, it is based on various ML models and human hearing abilities. Also, wa Maina [27] proposed to utilizes state-of-the-art advances in ML and signal processing to perform census in a joint framework, bird species detection, and biodiversity monitoring.

4.4 ML for Forest Fire Detection and Prevention

The task of detecting and predicting is quite challenging as they are so uncertain as they ignite and spreads very quickly. However, with the help of AI and IoT it's possible to create an early-warning systems with good accuracy by using some sensors that captures or go through the temperature, wind speed, humidity and it also detects the presence of smoke and fire. In some related studies, Thakur et al. [28] proposed the model that merges different data types, like genomic, clinical, and lifestyle factors, to make a precise risk assessments for individuals. By leveraging ensemble learning techniques, our framework delivers remarkable enhancements in both reliability and accuracy as compared to the pre-existing models, Toyin et al. [29] presented a widely used network intrusion detection system using ML model to capture and analyse network traffic. While, Giri et al. [30] proposed a DL model for cervical cancer detection using high resolution images that uses three distinct feature selection models. Similarly, AlMasri et al. [31] presented an idea where he used methods to defend and protect the network against the denial of services and attacks like port scanning. Also, Alam et al. [32] presented a fraudulent detection model that is very efficient with accuracy of 99.62%.

4.5 ML for Monitoring Illegal Logging Activities

The term Illegal logging [7], Here UAVs along with IoT can be used with AI. These old methods can be overcome, Manimozhi et al. [33] highlighted an AI driven nutrition monitoring. In another work, Kalita et al. [34] worked upon a cloud-based remote patient monitoring system that covers continuous glucose monitoring (CGM) with physical activities tracking to improve diabetes management using ML algorithms to deal problems in real-time with platform like Google Cloud Platform (GCP). Similarly, Chaudhry et al. [35] proposed using RF and SVM algorithms in ML to explore the applications, implementation, and assessment of the health parameters in the elderly healthcare sector as this algorithm has high accuracy of prediction and health monitoring devices that analyse heart rate, fall risks, and hypertension etc in real time. Also, Mishra and Singh [36] proposed an idea using advanced ML models to make a model which accurately analyses and predicts the fitness score of an individual based on their health parameters, Ramu et al. [37] presented an approach decentralised air quality monitoring that uses Edge AI and powerful ML techniques fitted with IoT sensors to collect air quality data using NVIDIA Jetson devices.

5 CONCLUSION

ML has started to be an integral part of scientific investigation into environmental issues, with deforestation a problem that still threatens natural habitats and disrupts climate conditions affecting societies at large all over the globe seen as one of today's most important areas of application. In this paper we reviewed the main ML techniques used in this field and how they contribute to detecting shifts of vegetation, studying satellite observations and increasing the precision of large-scale forest monitoring. But there remain several challenges that hinder the more extensive application of such approaches, e.g., regional forest peculiarity is inconsistent, lack of reliable labelled datasets and the computation cost of advanced models. One of the limitations is that several interesting research lines are pointed out, for example, developing models adjusted to certain geographic settings or considering different remote sensing data source as well as recent algorithms capable of near real time processing. In summary, ML offers strong potential for improving deforestation monitoring and supporting sustainable environmental planning, but continued work is needed to overcome the practical and technical barriers that remain.

References

[1] P. de Bem, O. de Carvalho Junior, R. Fontes Guimaraes, and R. Trancoso˜ Gomes, "Change detection of deforestation in the Brazilian amazon using Landsat data and Convolutional Neural Networks," Remote Sensing, vol. 12, no. 6, p. 901, 2020.

[2] I. Md Jelas, M. A. Zulkifley, M. Abdullah, and M. Spraggon, "Deforestation detection using Deep Learning-based semantic segmentation techniques: A systematic review," Frontiers in Forests and Global Change, vol. 7, 2024.

[3] S. Saha, S. Bhattacharjee, P. K. Shit, N. Sengupta, and B. Bera, "Deforestation probability assessment using integrated machine learning algorithms of Eastern Himalayan Foothills (India)," Resources, Conservation & Recycling Advances, vol. 14, p. 200077, 2022.

[4] Y. Tarazona, A. Zabala, X. Pons, A. Broquetas, J. Nowosad, and H. A. Zurqani, "Fusing Landsat and SAR data for mapping tropical deforestation through machine learning classification and the pvts-nonseasonal detection approach," Canadian Journal of Remote Sensing, vol. 47, no. 5, pp. 677–696, 2021.

[5] F. Pendrill et al., "Disentangling the numbers behind agriculture-driven tropical deforestation," Science, vol. 377, no. 6611, 2022.

[6] M. Kasahun, "Quantifying deforestation drivers through multi-temporal LULC analysis and population-forest correlation modeling: A case study of dara woreda, Ethiopia," Environmental Challenges, vol. 19, p. 101163, 2025.

[7] I. Carry and G. Maihold, "Illegal logging, timber laundering and the global illegal timber trade," in Geopolitics of the Illicit, 2022, pp. 275–308.

[8] S. Srivastava and T. Ahmed, "DLCD: Deep learning-based change detection approach to monitor deforestation," Signal, Image and Video Processing, vol. 18, no. S1, pp. 167–181, 2024.

[9] G. Aziz, N. Minallah, A. Saeed, J. Frnda, and W. Khan, "Remote Sensing based forest cover classification using Machine Learning," Scientific Reports, vol. 14, no. 1, 2024.

[10] I. H. Sarker, "Machine learning: Algorithms, real-world applications and Research Directions," SN Computer Science, vol. 2, no. 3, 2021.

[11] R. V. Maretto, L. M. Fonseca, N. Jacobs, T. S. Korting, H. N. Bendini, and L. L. Parente, "Spatio-temporal deep learning approach to map deforestation in Amazon Rainforest," IEEE Geoscience and Remote Sensing Letters, vol. 18, no. 5, pp. 771–775, 2021.

[12] F. AlMahamid and K. Grolinger, "Reinforcement learning algorithms: An overview and classification," in 2021 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), 2021, pp. 1–7.

[13] Z. Wang et al., "A deep learning based platform for remote sensing images change detection integrating crowdsourcing and active learning," Sensors, vol. 24, no. 5, p. 1509, 2024.

[14] Z. E. Sofi, I. Aouraghe, G. Khaissidi, M. Mrabti, and B. Faouzi, "Early detection of neurodegenerative diseases based on Moroccan population gait analysis," in 2024 International Conference on Connected Innovation and Technology (ICCITX), 2024, pp. 1–6.

[15] N. P. Chintya, G.-S. Jeon, S. Baek, and W. Kim, "Developing deep learning based change detection model for breakwaters observed by UAV (unmanned aerial vehicle)," Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography, vol. 40, no. 6, pp. 595–602, 2022.

[16] J. Serra et al., "Validation of AI models for ITCZ detection from Climate Data," in 2022 5th International Conference on Data Science and Information Technology (DSIT), 2022, pp. 1–7.

[17] S. Li, A. Pandey, B. Hooi, C. Faloutsos, and L. Pileggi, "Dynamic graphbased anomaly detection in the electrical grid," IEEE Transactions on Power Systems, vol. 37, no. 5, pp. 3408–3422, 2022.

[18] B. Shayesteh, C. Fu, A. Ebrahimzadeh, and R. H. Glitho, “Adaptive feature selection for predicting application performance degradation in Edge Cloud Environments,” IEEE Transactions on Network and Service Management, vol. 22, no. 1, pp. 121–138, 2025.
[19] M. V N, H. N, J. M, J. S. Reddy, and S. K. K, “Integrating ensemble and classical ML models for predicting CKD,” in 2024 Asia Pacific Conference on Innovation in Technology (APCIT), 2024, pp. 1–6.
[20] A. Rampal, N. Kapoor, and R. Saini, “Hybridization of Classification ML algorithms to predict alzheimer disease,” in 2024 2nd World Conference on Communication & Computing (WCONF), 2024, pp. 1–5.
[21] V. Chowdary B, C. Datta M, and R. Senapati, “An improved cardiovascular disease prediction model using ensembling of diverse machine learning classifiers,” in 2023 OITS International Conference on Information Technology (OCIT), 2023, pp. 329–333.
[22] Y. He, Y. Xiong, and Y. Tsai, “Machine learning based approaches to predict customer churn for an insurance company,” in 2020 Systems and Information Engineering Design Symposium (SIEDS), 2020, pp. 1–6.
[23] M. Molina-Rotger, A. Moran, M.´ A. Miranda, M. Roca, and B. Alorda-´ Ladaria, “BioVIEW: Automatic Remote Platform for Biodiversity Evaluation Using Low Cost Electronics,” in 2024 IEEE 2nd Conference on AgriFood Electronics (CAFE), 2024, pp. 105–109.
[24] L. Lang et al., “Autonomous bat echolocation monitoring system with machine learning for real-time Pest and biodiversity management,” in 2024 International Conference on Decision Aid Sciences and Applications (DASA), 2024, pp. 1–5.
[25] R.-T. Su et al., “Remote sensing classification study of grassland degradation in Inner Mongolia based on machine learning,” in IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium, 2024, pp. 10088–10091.
[26] K. Lopez-de-Ipina, M. Iturrate, J. B. Alonso, and B. Rodriguez-Herrera, “Automatic Acoustic Analysis for Biodiversity Preservation: A multienvironmental approach,” in 2015 4th International Work Conference on Bioinspired Intelligence (IWOBI), 2015, pp. 43–48.
[27] C. wa Maina, “BIOACOUSTIC approaches to Biodiversity Monitoring and conservation in Kenya,” in 2015 IST-Africa Conference, 2015, pp. 1–8.
[28] G. K. Thakur, N. Khan, H. Anush, and A. Thakur, “Ai-driven predictive models for early disease detection and prevention,” in 2024 International Conference on Knowledge Engineering and Communication Systems (ICKECS), 2024, pp. 1–6.
[29] O. Toyin et al., “Intelligent Network Intrusion Detection and Prevention System (NIDPS): A Machine Learning and network security,” in 2024 IEEE 5th International Conference on Electro-Computing Technologies for Humanity (NIGERCON), 2024, pp. 1–6.
[30] T. Giri, V. P. Meena, V. Sharma, K. Sharma, and P. Pawan, “Towards precision medicine: Deep learning models for Cervical Cancer Detection and Prevention,” in 2024 1st International Conference on Advanced Computing and Emerging Technologies (ACET), 2024, pp. 1–6.
[31] T. AlMasri, M. A. Snober, and Q. A. Al-Haija, “IDPs-SDN-ML: An intrusion detection and prevention system using software-defined networks and machine learning,” in 2022 1st International Conference on Smart Technology, Applied Informatics, and Engineering (APICS), 2022.
[32] M. S. Alam, R. K. Tiwari, and V. Pandey, “Healthcare Billing Fraud Detection through machine learning and using homographic encryption technique for prevention,” in 2021 4th International Conference on Recent Trends in Computer Science and Technology (ICRTCST), 2022, pp. 181–186.

[33] I. Manimozhi, N. G. N, B. L. B, R. Kanchana, and N. Divyashree, “Nutrition Reinvented: AI-Powered System For Personalized Dietary Insight,” International Journal of Environmental Sciences, vol. 11, no. 20s, pp. 1003–1014, 2025.
[34] D. Kalita, H. Sharma, and K. B. Mirza, “Cloud-based real-time remote patient monitoring system with incorporation of physical activity in closed loop diabetes management,” in 2024 IEEE Region 10 Symposium (TENSYMP), 2024, pp. 1–6.
[35] S. P. Chaudhry, A. Malode, and O. V. G. Swathika, “ML based multifaceted health monitoring devices: Enhancing elderly well-being,” in 2024 2nd International Conference on Recent Trends in Microelectronics, Automation, Computing and Communications Systems (ICMACC), 2024, pp. 526–531.
[36] S. Mishra and A. Singh, “IOT ML driven holistic health monitoring and fitness assessment empowering proactive wellbeing management,” in 2024 IEEE Students Conference on Engineering and Systems (SCES), 2024, pp. 1–6.
[37] V. Ramu et al., “Real-time air quality monitoring with edge AI and machine learning algorithm,” in 2024 8th International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2024, pp. 1204–1209.