International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Spectrum Allocation in Cognitive Radio based Traffic Monitoring System Using Machine Learning

Dr. N. Suganthi*[1], Dr. Suresh Kumar. K [2], Dr. Karthi Govindharaju[3]

Submitted:10/03/2024 **Revised**: 25/04/2024 **Accepted**: 02/05/2024

Abstract: Vehicle tracking and Traffic Monitoring is essential as it forms the main dimension of a smart city. Globally, during the last decade the number of automobiles in roadways has increased drastically. Traffic monitoring in such a high traffic density era is significantly difficult in various developing countries. Hence, the work focuses on regulating traffic jams by tracking the vehicle and transmitting the data to the regulating authorities in shorter duration with the help of Cognitive Radio technology. The CR technology is very useful for effective traffic monitoring to transmit the traffic management parameters by exploiting Primary User's (PU) spectrum. For spectrum detection and allocation for high-speed transmission of traffic parameters, various tree related machine learning algorithms like random forest, decision trees and XGBoost are used, examined and compared for better results. Of these, random forest gives high accurate prediction of available spectrum and allocation. On applying the model, we ensure that timely delivery of traffic monitoring information can help in better traffic management and vehicle tracking.

Keywords: *vehicle tracking, traffic management, cognitive radio networks, spectrum allocation*

1. Introduction

To enable Intelligent transport system applications, traffic monitoring and vehicle tracking are key factors. With urbanization of cities, the population density grows rapidly because of which the traffic density also grows rapidly. Intelligent traffic systems should allow the maximum traffic in the road network but at the same time should respond immediately to the road network traffic in order to avoid road congestion. The contemporary smart city scenario is facilitated by the growing evolution and the use of the Internet of Things (IoT) and th Internet of Everything (IoE) technologies, which are driving the smart city archetype to the big data scale.

Developing an automated and intelligent city system requires taking municipal data and turning it into actionable knowledge or insights, then creating a corresponding data-driven model [1]. Solutions for smart cities concentrate on a number of areas, including intelligent transportation, industry, energy, environment, health care, and living and infrastructure [2]. The main tenets of smart cities, including smart life, smart government, smart communications and smart environment, is the intelligent transport system [3]. Millions of drones, unmanned aerial vehicles (UAV) and autonomous cars will be a safe and environmentally friendly mode of transportation by 2030. To ensure the safety of passengers and pedestrians, they must meet strict requirements for localization, latency, and reliability. Traffic monitoring systems can help many applications including shortest route detection, automatic parking system, accident prone area, high traffic route for timely healthcare and more and more [4].

The number of video surveillance systems in cities is currently increasing quickly. These systems consist of different resolution and fixed frame rate video cameras as well as different mounting locations and resolutions [5]. Vehicles can serve as performance indicators for the transportation system if quantitative and qualitative road traffic characteristics are continuously monitored from fixed cameras. Low counting accuracy, categorizing a limited number of vehicle types, and tracking an object while determining its speed and direction of travel in all directions as it crosses the intersection's functional zone are the most commonly noted issues when evaluating real-time data from street cameras. Despite the obvious advantages of setting up such systems, not much study has been done to gather and analyze the speed and gesticulation patterns of traffic flows utilizing survey street cameras [6].

Traffic Monitoring Systems can be well organized by enabling the traffic data to be circulated to data centers in a quicker and faster way with the help of Cognitive Radio Networks (CRN). Since traffic monitoring involves real time streaming of video surveillance CRN will be the best option to communicate real time traffic.

Cognitive Radio Networks (CRNs) clout the alterable mechanism to efficaciously utilize underutilized wireless spectrum abundantly available. In recent decades, a diverse spectrum of interests has been drawn to CRNs. A CR

[1] *Assistant Professor, Department of Computer Science and Engineering, SRM IST, Ramapuram, Chennai, Tamil Nadu, India*
ORCID ID: 0000-0001-7022-6898
[2]*Professor & Head, Department of Information Technology, Saveetha Engineering College, Chennai, Tamil Nadu, India*
ORCID ID: 0000-0002-0673-3331
[3] *Associate Professor & Head, Department of AIDS, Saveetha Engineering College, Chennai, Tamil Nadu, India*
ORCID ID: 0000-0003-2486-237X
** Corresponding Author Email: suganthn@srmist.edu.in*

network has two users namely Primary User (PU) also called as licensed user and the second user is Secondary User (SU) also called as Unlicensed user. One of their most alluring features is that they may allow a large number of unlicensed users, often referred to as Cognitive Users (CUs), to use licensed spectrum bands opportunistically, which will help to maximize the use of a licensed spectrum more effectively [7,8]. The CRNs routing protocol has drawn a lot of attention recently, since it uses the idle licensed spectrum to send and receive data packets between the two CUs. This provides the desired benefits for many potential applications, including IIoT, smart cities, the military, healthcare, and self-driving cars. For instance, this protocol will improve the send/receive spectrum for data packets in smart city applications, improve smart application communication, and lower user costs [9].

With rapid development in smart cities, intelligent transportation systems can incorporate Cognitive Radio technology to address the main issue of band limited spectrum to communicate real time information about road traffic [10,11].

Spectrum prediction is an important venture in CR networks where SUs senses the spectrum availability based on the absence of PU. Because spectrum prediction necessitates the ongoing monitoring of variables like transmission rate and channel status, it is not only a difficult task but also one that is computationally expensive to accomplish.

In recent days, Machine learning algorithms are used to predict the availability of channels for transmitting real time data. To improve the accuracy of spectrum prediction, tree-based classifier models are used for intelligent transportation systems. Researchers have used different machine learning algorithms for traffic detection by vehicle detection. The main aim of our work is to use Cognitive Radio technology for fast communication of traffic related information among the users in CR Network. For spectrum sensing in the CR network, various tree-based classifier models are used for comparison.

2. Literature Survey

2.1. Traffic Management

Traffic control can be achieved through the use of tracking and video surveillance [12]. The construction of an intelligent transportation system is aided by the important information that surveillance cameras provide, such as traffic density and vehicle information. Manual traffic monitoring and analysis is a laborious task. Road traffic density needs to be measured, particularly in large cities, in order to properly manage traffic and operate traffic signals.

In [13,14], the authors propose an approach to traffic density estimation that does not include vehicle tracking. Four distinct picture thresholding methods [14] were used with the expectation–maximization (EM) algorithm [13] to determine which was the most accurate. Determining the image threshold might be considered an extreme form of contrast enhancement, i.e., making dark pixels darker and brilliant pixels lighter, for the purposes of image-in-video extraction and recognition. An intelligent transportation system's primary foundation is localization. Intelligent transportation relies heavily on the precise and quick localization of vehicles. Generally, nodes and automobiles rely on GPS signals for localization because of the highly dynamic nature of the vehicular network, vehicle mobility, and signal weakness caused by unfavourable wireless channel conditions.

In rough terrain, GPS accuracy is only 10 meters or greater [15]. The accuracy of GPS-based localization has recently been increased through the use of cooperative localization. A distributed cooperative localization technique was presented by the authors in [16] that fuses V2X measurements through particle filtering. This technique is used in tunnels. Deep neural networks and low-resolution video surveillance system data are used in some research [17] to measure traffic density and count the number of autos on the route. Conventional machine vision techniques are exemplified by the systems created in [18], which examined freight traffic issues. The majority of recent works address the modification and enhancement of contemporary detection systems, including SSD [19], Yolo [20], and Faster R-CNN and genetic algorithms [21], in order to identify a vehicle. The current approaches to real-time vehicle detection and classification impose stringent installation location and camera performance requirements in addition to requiring substantial processing power.

Most research works focussed on video surveillance and localization of vehicles which is highly hectic on implementation.

2.2 CRN and Machine Learning Techniques

A Cognitive radio network is an insightful and intuitive device that focuses on efficient channel utilization [22]. A CRN, or cognitive radio network, is a wireless system that uses resource allocation to manage current radio spectrum usage. The proliferation of wireless and mobile devices has led to problems with spectrum availability and resource allocation. CRNs have shown to be a very effective remedy for this issue [23]. The topic of spectrum assignment and access for CRNs is covered in [24], with an emphasis on interference that affects both users. In [25], a power mixture strategy for spectrum allocation is given. This technique increases the through-put to a CRN while satisfying the interference requirements for both users. The invasive weed optimization approach is suggested in [26] in order to improve the spectrum handoff efficiency, which represents load balancing and lowers handoff latency. Spectrum resources are allocated to smart grid users equitably in [27] using the standard grid configuration (SGCN). For computational reasons related to mathematical structure, wireless multiple-access channels have been utilized for PUs and the optimization of the energy efficiency (EE) problem seen in cognitive systems [28].

Using cognitive radio and machine learning, a smart and intelligent traffic system can be created that is both successful and efficient [29,30]. Spectral and energy-efficient smart traffic management systems and devices are necessary when they work with cognitive radiology [31,32]. The goal of machine learning, a well-known branch of computer science, is to create algorithms and software that can be tested and trained on a variety of interesting data sets and that can respond intelligently to new information [33]. The development of contemporary technologies, including speech recognition, computer vision, image processing, and object (face, text, posture, and people) detection in robots, depends on machine learning approaches and algorithms [34, 35]. The intelligent traffic system can benefit from machine learning (ML) to convert traditional traffic management systems into smart and intelligent, given the successful application of ML in other scientific and technological domains. By integrating ML and CR technologies, it will be possible to realize the vision of intelligent traffic systems by enabling sensing and monitoring equipment to adjust their parameters in real-time for data transmission and processing based on dynamic radio conditions.

In CRN, the main key feature where Machine learning algorithms can be applied are Spectrum Sensing (SS). One of the most predominant functions of a CR is spectrum sensing, which determines whether a PU is present or not. The SS problem can be stated as a classification problem, with a binary value as the output that indicates channel availability. For Supervised ML algorithms, let us assume M number of training data points a = {a (1), a (2),….. a(M)} with corresponding output labels b = {b (1), b (2), …… b(M)} that indicates the channel availability or unavailability. When an unknown data point a(i) is provided to the trained model, the model has to classify it to one of the class b(i), b(i) = 1 channel availability and b(i) = 0 channel unavailability. Support Vector Machine (SVM), k closest neighbour (KNN), Bayesian classifier, and artificial neural networks (ANN) are the most often utilized supervised machine learning algorithms in SS.

In [36], every secondary user provides sensory data to the fusion centre, which analyses this data to determine whether the principal user is present worldwide. Before using the combined signal to obtain a local result, a selection combiner (SC) combines the Energy Detector (ED) outputs with signals from the primary user (PU), which are established by various antennas on SU. A hybrid Support Vector Machine (SVM) is used at the Fusion Center (FC) to expunge SUs, which greatly enhances detection performance and lowers the quantity of false positives. In order to improve SVM performance in terms of detection probability and misclassification risk, two phase SVM is used in [37]. The classifier is trained using the energy levels of the PUs as feature vectors. A high-dimensional feature space was first created by mapping the incoming signal, and then SVM was employed for further classification in [38] feature-based testing. Decision statistics like energy detection, maximum-minimum eigenvalue ratios, and their higher order combinations make up the feature vector.

In order to increase spectrum sensing's identification rate, the paper [39] suggested using machine learning (ML) to optimize the RBF method. RBF algorithm is used in conjunction with SVM. The SVM/RBF method outperformed the RBF approach in terms of average spectrum detection success rate. This suggests that using machine learning to analyze the RBF neural network technique can increase spectrum sensing's success rate. The paper [40] proposes an effective feature extraction and reduction method-based SS model for CR based on machine learning (ML). The five stages of the proposed study are featuring extraction, dimensionality reduction, wavelet transform, noise removal, and classification. Ensemble machine learning classifiers such as Support Vector Machine (SVM), Naive Bayes (NB), and K-Nearest Neighbour (KNN) are used to identify whether the PU signal is active or not. To evaluate the effectiveness of the models for SS that are presented, simulations are run. The outcomes demonstrated that SVM, with its higher accuracy and lower SNR, achieves the greatest performance for SS.

Through their parallel connection, long-short-term memory (LSTM) and convolutional neural networks (CNN) provide complementary feature extraction capabilities that are completely utilized by the cooperative spectrum sensing model developed by the study [41]. Among them, the CNN is used to extract hidden spatial information, and the LSTM network is used to extract time characteristics. When the network is connected serially, CNN and LSTM can both process the original dataset directly, preventing information feature loss. Three important CR-VANETs concerns are presented in this research [42]: optimal channel allocation to CR users, channel indexing for selective SS, and dependable Cooperative Spectrum Sensing (CSS). All three are addressed in a single framework. In CSS, Deep Reinforcement Learning (DRL) technique is applied to achieve the global CSS session by combining the local SS choice with more crucial features like the timestamp and the location of the sensing signal acquisition. To reduce CR users' sense overload, selective channel-based spectrum sensing is crucial. The Long Short-Term Memory (LSTM) model, which is based on deep learning, is used in this study's time series analysis to index the key user channels for selective SS. In the end, we formulate the complicated environment as a Partial Observable Markov Decision Process (POMDP) framework and solve the channel allocation to the CR-VANETs through a value iteration approach.

Observations from Literature survey

- In most of the research works, SVM, KNN, ANN and deep learning algorithms are used for SS in CRN.
- The feature used in most of the ML and DL algorithms is Energy Statistics.

Hence, the proposed model in this research is carried out with the help of tree-based machine learning algorithms like decision trees, random forest and XGBoost with probability vector features. Finally, the model is applied in real time traffic management and serves as a support for Intelligent Transportation Systems.

3. System Model

Figure 1 illustrates how the cognitive radio in the intelligent transportation system operates. To monitor the traffic conditions in various regions of the city, various sensing and monitoring nodes are installed on the traffic road. These nodes also conduct spectrum sensing at specific time slots, sharing the monitored data with the intelligent transportation system's central control.

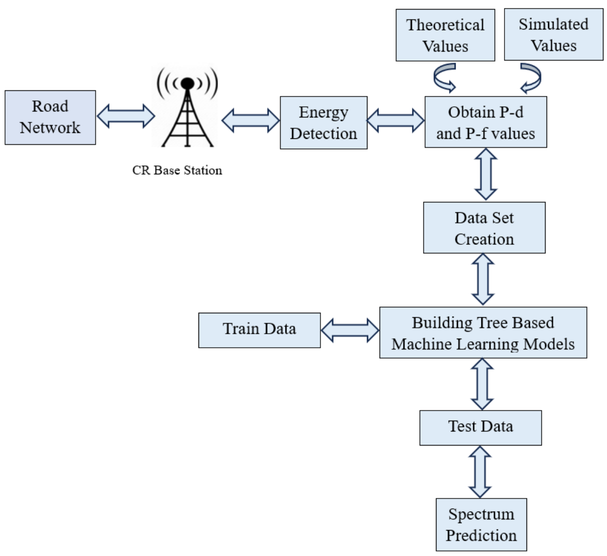

Fig 1. Cognitive Radio and tree-based ML models in Spectrum Prediction

The same time frame (T) is shared by all sensor nodes, and it is further split into two time slots: T-s, or sensing, and T-tr, or transmission. When the principal users' spectrums are judged free for use, the slot-sensing nodes in T-s transmit their data to the fusion centre or CR base station in the T-tr slot. To exchange the data gathered from the sensing nodes, however, the CR base station speaks with other parts of the intelligent transportation system. Secondary users' sensed data is classified using tree-based algorithms (TBAs). The spectrum sensing's training and testing accuracy is assessed to determine whether or not the spectrum is open for use. Let us consider an intelligent transportation system with a centralized CR-based Sensor Network that comprises M wireless sensing nodes, further referred to as secondary users (SUs). The SUs can conduct spectrum sensing by applying a binary hypothesis test to the received signal. This can be stated as

$$Signal\ (n) = \begin{cases} Primary(n) + Noise(n), & Hyp(1) \\ Noise(n), & Hyp(0) \end{cases} \quad (1)$$

where Primary(n) represents the primary user's signal and Noise(n) represents Additive White Gaussian Noise (AWGN). The binary hypothesis Hyp(0) and Hyp(1) represents the absence and presence of Primary User, respectively. There are other methods for doing spectrum sensing, but energy detection is the most widely used since it is simple to use and doesn't require any knowledge of the primary signal.

Process demonstrates that SUs continuously scans the surrounding environment to gather the N number of received signal samples in accordance with Equation (1). The average energy of these N received samples can be calculated by taking the square of the magnitude of each sample and averaging the total number of received samples. It is necessary to compute a preset detection threshold in order to achieve a high target detection probability. After that, a preset detection threshold is compared to the average energy of the samples that were received. The final stage involves making a conclusion based on the comparison between the detection threshold and the average energy of the signal. The principal user is recognized as being present in the measured spectrum if the average energy exceeds the detection threshold; otherwise, SUs is free to use the spectrum. The average energy of the received primary signal samples, or test statistic T, is compared to the detection threshold. T can be written as

$$T = \frac{1}{N} \sum_{1}^{N} (Signal(n))^2 \quad (2)$$

Two crucial parameters related to spectrum sensing are probabilities of detection (P-d) and false alarm (P-f), which also demonstrate the effectiveness of CRNs. High (P-d) and low (P-f) are always necessary for an effective CRN. While (P-f) provides the incorrect probability of PU's presence in the provided spectrum, (P-d) truly provides us with the likelihood of PU's presence in the supplied spectrum. As a result, high (P-d) is always necessary to prevent SUs from interfering with PUs. However, low (P-f) is required for effective CRNs since it represents a lost chance to utilize the open spectrum. By comparing T with the established detection threshold λ, one can often calculate (P-d) and (P-f) based on whether the primary users are present in the spectrum. P-d and P-f can be defined as [43] in terms of T and λ.

$$P - d = \Pr(T > \lambda \mid Hyp(1) \quad (3)$$

$$P - f = \Pr(T > \lambda \mid Hyp(0)) \quad (4)$$

The first step in the spectrum sensing procedure is to divide the received signal into N samples, depending on the needs. Both (P-d) and (P-f) rely on N. Higher (P-d) and lower (P-f) are guaranteed by larger values of N. Thus, (P-d) and (P-f) have the following expressions [44]:

$$P - d = Q\left(\left(\frac{\lambda}{\sigma_n^2} - \gamma - 1\right)\sqrt{\frac{N}{2\gamma+1}}\right) \quad (5)$$

$$P - f = Q\left(\left(\frac{\lambda}{\sigma_n^2} - 1\right)\sqrt{N}\right) \quad (6)$$

where Q(.) is called the Q-function, defined by

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} exp\,(-\frac{v^2}{2})\, dv \quad (7)$$

To reduce the likelihood of missing PU detections, a greater target detection probability (P-d1) is assumed in spectrum sensing. Interference with PU communication results from the failure to detect PU, which is undesired in CRNs. Consequently, one can utilize (P-d1) to get such a value of λ, which can be useful to get both lower (P-f) and higher (P-d):

$$\lambda = \sigma_n^2(\gamma + 1 + \sqrt{\frac{2\gamma+1}{N}}\, Q^{-1}(P - d1) \quad (8)$$

Steps:

1. Energy detector: The entire process of energy detection involves comparing the received signal to the average energy of received signal samples using a pre-defined detection threshold.
2. Calculating P-d and P-f: Theoretical and simulated values of P-d and P-f are calculated based on the equation 5, 6 and 7.
3. Data Set creation: To construct a set of values for two probabilities, the estimated values of P-d and P-f are supplied into the data generation block.
4. Data training: During the data-training phase, the model is trained for 70% of the values in the data set.
5. Data testing: 30% of the data set is left over after the model has been trained and tested.
6. Evaluation by various tree-based ML algorithms

4. Evaluation Metrics

A variety of metrics were used to evaluate the effectiveness of a model and the generalization ability of the trained classifier. In order to find the optimal classifier during classification training, the assessment metrics are essential.

4.1 Precision

This is one of the evaluation measures that indicates the ratio of correctly anticipated positive outcomes to all expected positive values. A different way to define this would be a measure of the percentage of correctly predicted positive patterns in a positive class compared to all of the predicted patterns [51].

$$Precision = \frac{TP}{TP+FP} \quad (9)$$

4.2 Accuracy

By calculating the ratio of correctly predicted positive and negative values to the total number of assessed cases, this metric evaluates the overall performance of the model [51].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (10)$$

5. Results and Discussions

5.1. Data Modeling

The spectrum sensing technique yields two distinct data sets, one derived from the theoretical approach and the other from the simulation process.

Equations 5 and 6 are used to get the theoretical values of the probability of detection (P-d) and the chance of false alarm (P-f). Results from Matlab simulations are used to assess how well various tree-based machine learning models perform. SNR () = -10dB, target detection probability (P-d1) = 0.8, and frequency are set to fs = 8MHz. Spectrum-sensing data evaluation is done using Matlab's analysis of the categorization learning tool [69]. We generate the random noise and primary signal samples in Matlab to build an environment. To produce the noise and main signal independently, 500 Monte Carlo simulations are used. Each simulation calculates the average energy of the received signal for both noise only and noise with primary signal. The predetermined energy detection threshold is used to compare the average energy of the two scenarios. Pf is the number of times the average energy exceeds the threshold under hypothesis H0 (only noise) following 1000 runs. The total number of simulations is then split by this figure. This process computes values of 100 Pf. Similarly, Pd is calculated as follows: under hypothesis H1 (noise + primary signal), Pd is equal to the total number of simulations divided by the number of times the average energy surpasses the threshold. The simulated values of Pd and Pf are computed in this manner. The classifiers are trained using 70% of the total data from both the theoretical and simulated data sets, with the remaining 30% being utilized for testing.

All tree-based models are compared in terms of training (validation) and testing accuracy after being trained and tested in the classification learner. Tables 1 and 2 offer an analysis of the testing and training accuracies of many tree-based classifiers using simulated data for 500, 1000, and 1500 samples of the received signal. Tables 3 and 4 display the precision score.

Table 1. Comparison of tree-based models – Accuracy for simulated data

Classifier	500 samples		1000 samples		1500 samples	
	Training Accuracy (%)	Testing Accuracy (%)	Training Accuracy (%)	Testing Accuracy (%)	Training Accuracy (%)	Testing Accuracy (%)
Decision Tree	84.43	85.12	91.54	90.34	95.67	95.34
Gradient Boost	85.00	84.13	89.14	88.65	93.48	92.59
XGBoost	86.38	85.27	90.42	91.17	95.31	94.28
Random Forest	87.34	86.91	92.42	92.19	96.48	96.12

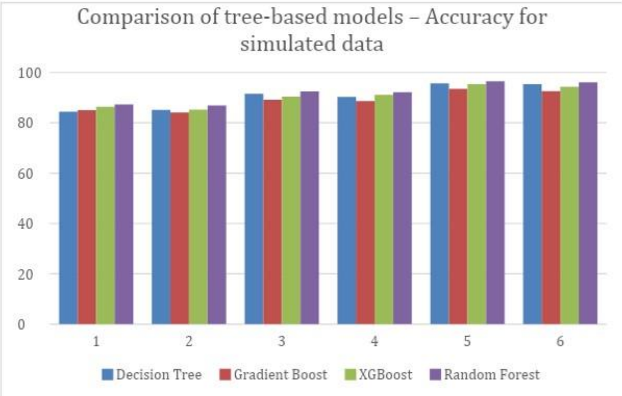

Fig 2: Accuracy Comparison for simulated data

The training and testing accuracies of several classifier types are shown in Table 1 and Figure 2. It is possible to deduce from the data that accuracy rises as sample count increases. The sensing findings alter in relation to P-d and P-f when the number of samples varies. In real-time scenarios, it exhibits high P-d and low P-f values as the number of samples grows. We assign a 1 to P-d and a 0 to P-f. It so gives the dataset more 1s than 0s. Consequently, classifiers that are taught with a large number of 1s yield high testing and training accuracies.

Table 2. Comparison of tree-based models – Accuracy for theoretical data

Classifier	500 samples		1000 samples		1500 samples	
	Training F1 score (%)	Testing F1 score (%)	Training F1 score (%)	Testing F1 score (%)	Training F1 score (%)	Testing F1 score (%)
Decision Tree	80.47	80.12	84.32	83.90	90.54	90.21
Gradient Boost	81.15	80.90	86.72	86.21	92.26	92.01
XGBoost	84.52	84.16	90.51	90.13	92.89	92.50
Random Forest	85.41	85.20	90.40	91.38	93.38	94.28

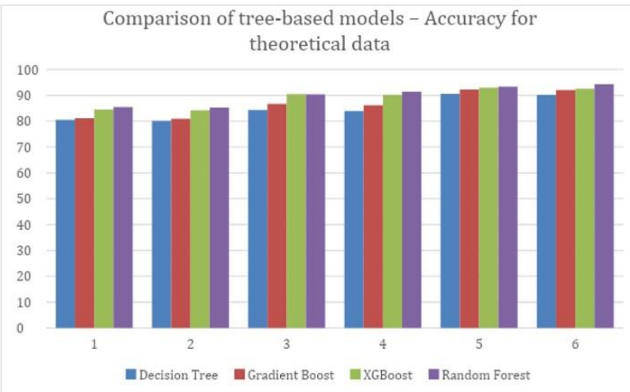

Fig 3: Accuracy Comparison for theoretical data

For the same theoretical data set, Table 2 compares the testing and training accuracies of each classifier for 500, 1000, and 1500 received signal samples.

Table 3. Comparison of tree-based models – Precision score for simulated data

Classifier	500 samples		1000 samples		1500 samples	
	Training Precision (%)	Testing Precision (%)	Training Precision (%)	Testing Precision (%)	Training Precision (%)	Testing Precision (%)
Decision Tree	79.85	78.65	84.39	83.25	91.40	90.11
Gradient Boost	80.54	80.20	85.35	84.29	92.12	91.89
XGBoost	83.16	82.32	87.72	88.63	92.73	93.34
Random Forest	88.28	88.04	90.38	91.12	93.57	94.10

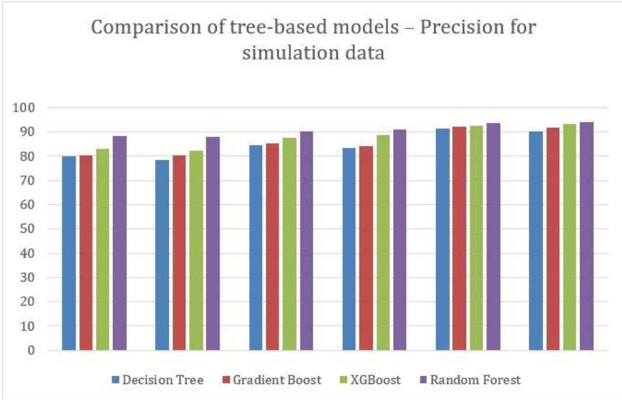

Fig 4: Precision Comparison for simulated data

Table 4. Comparison of tree-based models – Precision score for theoretical data

Classifier	500 samples		1000 samples		1500 samples	
	Training Recall (%)	Testing Recall (%)	Training Recall (%)	Testing Recall (%)	Training Recall (%)	Testing Recall (%)
Decision Tree	78.52	77.23	80.51	80.15	90.48	90.21
Gradient Boost	79.62	79.28	82.36	82.18	91.54	91.32
XGBoost	82.17	81.71	86.38	86.02	91.23	92.12
Random Forest	86.37	86.20	89.45	89.15	92.42	93.38

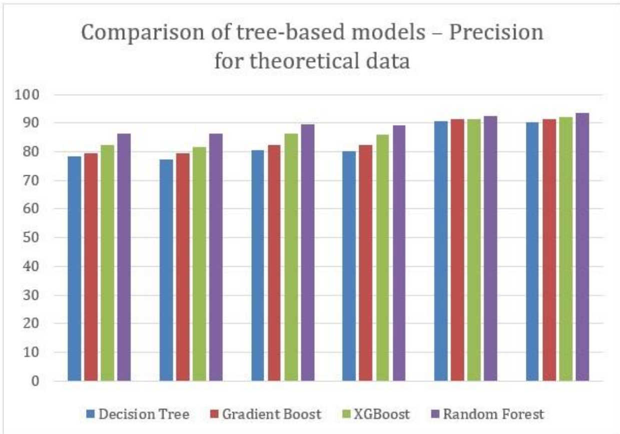

Fig 5: Precision Comparison for theoretical data

For the same simulated data set, Table 3 compares the precision scores of all classifiers for testing and training for 500, 1000, and 1500 received signal samples. For the same simulated data set, Table 4 compares the precision scores of all classifiers for testing and training for 500, 1000, and 1500 received signal samples.

The Random Forest classifier has a high accuracy and precision score based on the results of both theoretical and simulated data. The Random Forest tree classifier's theoretical and simulated data sets have average testing accuracies of 90.28% and 91.74%, respectively. Comparably, the Random Forest tree classifier's average testing precision scores for the theoretical and simulated data sets are 89.57% and 91.08%, respectively.

6. Conclusion

Smart Intelligent Transport systems are revolutionizing in recent research works where incorporating Machine learning algorithms, sensor devices, and spectrum utilization can modernize the transportation system. Cognitive Radio technology can be used to resolve spectrum scarcity problems for wireless sensor network-based applications. Machine learning algorithms help in efficient spectrum detection and prediction of available spectrum and helps CR technology for further efficient spectrum utilization. ML algorithms also help in receiving traffic related information to the base station via Primary user network without interference with other primary user communication. This paper uses tree-based machine learning methods to detect and forecast spectrum. The discovered energy value is utilized to forecast the spectrum's availability. Using both theoretical and simulation techniques, the data set for P-d and P-f is constructed using the detected energy values. The training and testing accuracy and precision scores are computed using several tree-based classifiers, such as random forest, gradient boost, XGBoost, and decision tree. The values of the simulated and theoretical data sets are used to apply the classifiers. The classifiers are used with varying sample ranges, beginning with 500, 1000, and 1500 samples. Random forest produces the highest accuracy and precision in testing and training out of all the classifiers. On theoretically produced values P-d and P-f data set, the random forest classifier's testing and training accuracies are 85.41% and 85.20% for 500 samples, 90.40% and 91.38% for 1000 samples, and 93.38% and 94.28% for 1500 samples. Using simulated P-d and P-f data sets, the random forest classifier's testing and training accuracies are 87.34% and 86.91% for 500 samples, 92.42% and 92.19% for 1000 samples, and 96.48% and 96.12% for 1500 samples.

Author contributions

Dr.N. Suganthi: Conceptualization, Methodology, Software, Writing-Original draft preparation **Dr. Suresh Kumar. K:** Writing-Original draft preparation, Software, Validation. **Dr. Karthi Govindharaju:** Conception, Examination, Writing-Reviewing and Editing.

Conflicts of interest

The authors declare no conflicts of interest.

Funding Information Not applicable

Data Availability Not applicable

Research Involving Human and /or Animals No

References

[1] R. Janani, K. Renuka, A. Aruna, et al., “Iot in smart cities: A contemporary survey,” Global Transitions Proceedings, vol. 2, no. 2, pp. 187–193, 2021.

[2] P. Bellini, P. Nesi, and G. Pantaleo, “Iot-enabled smart cities: A review of concepts, frameworks and key technologies,” Applied Sciences, vol. 12, no. 3, p. 1607, 2022.

[3] C. S. Lai, L. L. Lai, and Q. H. Lai, “Smart city,” in Smart Grids and Big Data Analytics for Smart Cities, pp. 1–171, Springer, 2021.

[4] M. Derawi, Y. Dalveren, and F. A. Cheikh, “Internet-of-Things-based smart transportation systems for safer roads,” in Proc. IEEE 6th World Forum Internet Things (WF-IoT), 2020, pp. 1–4.

[5] Peppa MV, Bell D, Komar T, Xiao W. Urban traffic flow analysis based on deep learning car detection from cctv image series. Int Arch Photogramm Remote Sens Spat Inf Sci. 2018;42(4):565–72. https://doi.org/10.5194/isprsarchives-XLII-4-499-2018.

[6] Fedorov A, Nikolskaia K, Ivanov S, Shepelev V, Minbaleev A. Traffic flow estimation with data from a video surveillance camera. J Big Data. 2019. https://doi.org/10.1186/s40537-019-0234-z.

[7] M.A. Rahman, A.T. Asyhari, M.Z.A. Bhuiyan, Q.M. Salih, K.Z.B. Zamli, L-CAQ: joint link-oriented channel-availability and channel-quality based channel selection for mobile cognitive radio networks J. Netw. Comput. Appl., 113 (2018), pp. 26-35

[8] S. Tyagi, S. Tanwar, N. Kumar, J.J. Rodrigues Cognitive radio-based clustering for opportunistic shared spectrum access to enhance lifetime of wireless sensor network Pervasive Mob. Comput., 22 (2015), pp. 90-112

[9] M. Rahim, M. Rahman, M.A. Rahman, A.J.M. Muzahid, S.F. Kamarulzaman, *et al.* A framework of iot-enabled vehicular noise intensity monitoring system for smart city International Conference on Innovative Technology, Engineering and Science, Springer (2020), pp. 194-205

[10] J. Si, H. Huang, Z. Li, B. Hao, and R. Gao, “Performance analysis of adaptive modulation in cognitive relay network with cooperative spectrum sensing,” IEEE Trans. Commun., vol. 62, no. 7, pp. 2198–2211, Jul. 2014.

[11] P. Goswami et al., “AI based energy efficient routing protocol for intelligent transportation system,” IEEE Trans. Intell. Transp. Syst., vol. 23, no. 2, pp. 1670–1679, Feb. 2022.

[12] Y. Liu, B. Tian, S. Chen, F. Zhu, and K. Wang,” A survey of vision-based vehicle detection and tracking techniques in its,”in Proc. IEEE Int. Conf. Veh. Electron. Saf., July 2020.

[13] N. Sammaknejad, Y. Zhao, and B. Huang, “A review of the expectation maximization algorithm in data-driven process identification,” Journal of Process Control, vol. 73, pp. 123-136, 2019.

[14] N. Senthilkumaran, and S. Vaithegi, S, “Image Segmentation by Using Thresholding Techniques for Medical Images,” Computer Science and Engineering: An International Journal, vol. 6, no. 1, 2016.

[15] N. Piperigkos, A. S. Lalos, K. Berberidis, and C. Anagnostopoulos, “Cooperative multi-modal localization in connected and autonomous vehicles,” in 2020 IEEE 3rd Connected and Automated Vehicles Symposium (CAVS), pp. 1–5, IEEE, 2020.

[16] G.-M. Hoang, B. Denis, J. Harri, and D. T. Slock, “Robust data fusion for ̈cooperative vehicular localization in tunnels,” in 2017 IEEE Intelligent Vehicles Symposium (IV), pp. 1372–1377, IEEE, 2017.

[17] Zhang F, Li C, Yang F. Vehicle detection in urban traffic surveillance images based on convolutional neural networks with feature concatenation. Sensors. 2019;19(3):594. https://doi.org/10.3390/s19030594.

[18] Rathore MM, Son H, Ahmad A, Paul A. Real-time video processing for traffic control in smart city using Hadoop ecosystem with GPUs. Soft Comput. 2018;22(5):1533–44. https://doi.org/10.1007/s00500-017-2942-7.

[19] Alipio, Melchizedek I., Kathlyn Mae T. Peñalosa, and Julioh Roscoe C. Unida. "In-store customer traffic and path monitoring in small-scale supermarket using UWB-based localization and SSD-based detection." *Journal of Ambient Intelligence and Humanized Computing* 14.5 (2023): 4955-4969.

[20] Mittal, Usha, Priyanka Chawla, and Rajeev Tiwari. "EnsembleNet: A hybrid approach for vehicle detection and estimation of traffic density based on faster R-CNN and YOLO models." *Neural Computing and Applications* 35.6 (2023): 4755-4774.

[21] Ghasemi Darehnaei, Zeinab, et al. "Ensemble deep learning using faster r-cnn and genetic algorithm for vehicle detection in uav images." *IETE Journal of Research* 69.8 (2023): 5102-5111.

[22] Lin, Z.; An, K.; Niu, H.; Hu, Y.; Chatzinotas, S.; Zheng, G.; Wang, J. SLNR-based secure energy efficient beamforming in multibeam satellite systems. IEEE Trans. Aerosp. Electron. Syst. 2022.

[23] Chuang, C.L.; Chiu, W.Y.; Chuang, Y.C. Dynamic multiobjective approach for power and spectrum allocation in cognitive radio networks. IEEE Syst. J. 2021, 15, 5417–5428.

[24] Alhussien, N.; Gulliver, T.A. Joint Resource and Power Allocation for Clustered Cognitive M2M Communications Underlaying Cellular Networks. IEEE Trans. Veh. Technol. 2022, 71, 8548–8560.

[25] Anumandla, K.K.; Sabat, S.L.; Peesapati, R.; AV, P.; Dabbakuti, J.K.; Rout, R. Optimal spectrum and power allocation using evolutionary algorithms for cognitive radio networks. Internet Technol. Lett. 2021, 4, e207.

[26] [26] Sumathi, D.; Manivannan, S. Stochastic approach for channel selection in cognitive radio networks using optimization techniques. Telecommun. Syst. 2021, 76, 167–186.

[27] Kaur, A.; Kumar, K. Energy-efficient resource allocation in cognitive radio networks under cooperative multi-agent model-free reinforcement learning schemes. IEEE Trans. Netw. Serv. Manag. 2020, 17, 1337–1348.

[28] Ostovar, A.; Zikria, Y.B.; Kim, H.S.; Ali, R. Optimization of resource allocation model with energy-efficient cooperative sensing in green cognitive radio networks. IEEE Access 2020, 8, 141594–141610.

[29] Aslam, Muhammad Muzamil, et al. "Beyond6G-consensus traffic management in CRN, applications, architecture and key challenges." *2021 IEEE 11th International Conference on Electronics Information and Emergency Communication (ICEIEC) 2021 IEEE 11th International Conference on Electronics Information and Emergency Communication (ICEIEC)*. IEEE, 2021.

[30] Ravish, Roopa, and Shanta Ranga Swamy. "Intelligent traffic management: A review of challenges, solutions, and future perspectives." *Transport and Telecommunication Journal* 22.2 (2021): 163-182.

[31] [Aslam, Muhammad Muzamil, et al. "Consensus performance of traffic management system for cognitive radio network: an agent control approach." *Cyberspace Data and Intelligence, and Cyber-Living, Syndrome, and Health: International 2019 Cyberspace Congress, CyberDI and CyberLife, Beijing, China, December 16–18, 2019, Proceedings, Part II 3*. Springer Singapore, 2019.

[32] Muhammad Muzamil Aslam, Liping Du, Xiaoyan Zhang, Yueyun Chen, Zahoor Ahmed, Bushra Qureshi, "Sixth Generation (6G) Cognitive Radio Network (CRN) Application, Requirements, Security Issues, and Key Challenges", Wireless Communications and Mobile Computing, vol. 2021, Article ID 1331428, 18 pages, 2021. https://doi.org/10.1155/2021/1331428

[33] Mitra, A.; Bera, B.; Das, A.K.; Jamal, S.S.; You, I. Impact on blockchain-based AI/ML-enabled big data analytics for cognitive Internet of Things

environment. Comput. Commun. 2023, 197, 173–185. [CrossRef]
[34] Xu, G.; Khan, A.S.; Moshayedi, A.J.; Zhang, X.; Shuxin, Y. The Object Detection, Perspective and Obstacles In Robotic: A Review. EAI Endorsed Trans. AI Robot. 2022, 1, e13. [CrossRef]
[35] Ragno, L.; Borboni, A.; Vannetti, F.; Amici, C.; Cusano, N. Application of Social Robots in Healthcare: Review on Characteristics, Requirements, Technical Solutions. Sensors 2023, 23, 6820. [CrossRef] [PubMed]
[36] Paramasivam Thuraipandi, Sivagurunathan and Nagarajan, Sathish Kumar. 'Cooperative Spectrum Sensing Based Hybrid Machine Learning Technique for Prediction of Secondary User in Cognitive Radio Networks'. 1 Jan. 2023: 3959 – 3971.
[37] Ghazizadeh E, Abbasi D, Nezamabadi-pour H., "An enhanced two-phase SVM algorithm for cooperative spectrum sensing in Cognitive Radio Networks", Int J Commun Syst 2018;32(2): e3856.
[38] Coluccia A, Fascista A, Ricci G., "Spectrum sensing by higher order SVM based detection", 27th European Signal Processing Conference (EUSIPCO). A Coruna, Spain; IEEE:2019.
[39] Yang, S., Tong, C. Cognitive spectrum sensing algorithm based on an RBF neural network and machine learning. Neural Comput & Applic 35, 25045–25055 (2023). https://doi.org/10.1007/s00521-023-08488-y
[40] Rose, Biji and Aruna Devi, B. 'Spectrum Sensing in Cognitive Radio Networks Using an Ensemble of Machine Learning Frameworks and Effective Feature Extraction'. 1 Jan. 2023: 10495 – 10509.
[41] Li, Liuwen, Wei Xie, and Xin Zhou. "Cooperative Spectrum Sensing Based on LSTM-CNN Combination Network in Cognitive Radio System." IEEE Access (2023).
[42] Paul, Anal, and Kwonhue Choi. "Deep learning-based selective spectrum sensing and allocation in cognitive vehicular radio networks." Vehicular Communications 41 (2023): 100606.
[43] Ali, M.; Nam, H. Optimization of spectrum utilization in cooperative spectrum sensing. Sensors 2019, 19, 1922. [CrossRef]
[44] Liang, Ying-Chang, et al. "Sensing-throughput tradeoff for cognitive radio networks." IEEE transactions on Wireless Communications 7.4 (2008): 1326-1337.
[45] Geron, A. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, 2nd ed.; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2019.
[46] Rokach, L.; Maimon, O., Decision Trees. In Data Mining and Knowledge Discovery Handbook; Maimon, O., Rokach, L., Eds.; Springer: Boston, MA, USA, 2005; pp. 165–192
[47] Dangeti, P. Statistics for Machine Learning, 1st ed.; Packt Publishing, Limited: Birmingham, AL, USA, 2017.
[48] Liu, Y.; Wang, Y.; Zhang, J. New Machine Learning Algorithm: Random Forest. In Proceedings of the Information Computing and Applications, Chengde, China, 14–16 September 2012; Liu, B., Ma, M., Chang, J., Eds.; Springer: Berlin/Heidelberg, Germany, 2012; pp. 246–252.
[49] Natekin, A.; Knoll, A. Gradient boosting machines, a tutorial. Front. Neurorobot. 2013, 7, 21.
[50] Dhieb, N.; Ghazzai, H.; Besbes, H.; Massoud, Y. Extreme Gradient Boosting Machine Learning Algorithm For Safe Auto Insurance Operations. In Proceedings of the 2019 IEEE International Conference on Vehicular Electronics and Safety (ICVES), Cairo, Egypt, 4 September 2019; pp. 1–5.
[51] Hossin, M.; Sulaiman, M.N. A Review on Evaluation Metrics for Data Classification Evaluations. Int. J. Data Min. Knowl. Manag. Process 2015, 5, 1–11.