Scientific Journal of King Faisal University:
Basic and Applied Sciences

A Machine Learning Framework for Spectrum Sensing and Occupancy Analysis Using Satellite Data

Sreerama Samartha J.G.[1], Archana N.V.[2], Dayananda GK[1] and Vayusutha M.[1]

[1] Department of Electronics & Communication Engineering, Canara Engineering College, Visvesvaraya Technological University, Mangalore, India.
[2] Department of Electronics & Communication Engineering, The National Institute of Engineering, Visvesvaraya Technological University, Mysore, India.

LINK	RECEIVED	ACCEPTED	PUBLISHED ONLINE	ASSIGNED TO AN ISSUE
https://doi.org/10.37575/b/eng/240050	20/11/2024	03/03/2025	03/03/2025	01/06/2025
NO. OF WORDS	**NO. OF PAGES**	**YEAR**	**VOLUME**	**ISSUE**
8432	10	2025	26	1

ABSTRACT

In satellite communication systems, effective spectrum sensing and management are essential, especially in scenarios involving both geostationary Earth orbit (GEO) and non-geostationary Earth orbit (NGEO) satellites. As the number of NGEO satellites increases, managing interference with GEO signals becomes more complex. This study introduces a machine learning (ML) framework for spectrum sensing, spectrum hole detection and occupancy prediction based on satellite data. The framework utilizes two ML models, support vector machine (SVM) and random forest (RF), along with a hybrid model combining both. SVM is used to classify spectrum occupancy based on GEO signals, while RF is employed to detect spectrum holes and predict future occupancy patterns. The hybrid model merges the strengths of both to enhance prediction accuracy and robustness. A comparative analysis of the models evaluated accuracy, computation time and robustness against interference. The results show that SVM achieved 99.17% accuracy, excelling in precision, while RF reached 99.12% accuracy, demonstrating better recall and more effective identification of occupied spectrum regions. The hybrid model outperformed both, achieving 99.25% accuracy, with an improved balance between precision and recall and superior performance under complex interference conditions. This study highlights the effectiveness of SVM, RF and their hybrid in optimizing spectrum management.

KEYWORDS
Hybrid Model, machine learning, random forest, satellite communication, spectrum management, support vector-machine

CITATION
Samartha J.G.S., Archana, N.V., Dayananda, G.K. and Vayusutha, M. (2025). A machine learning framework for spectrum sensing and occupancy analysis using satellite data. *Scientific Journal of King Faisal University: Basic and Applied Sciences*, **26**(1), 25–34. DOI: 10.37575/b/eng/240050

1. Introduction

Spectrum utilization is important in wireless communication mainly because of the growing demand. Both geostationary Earth orbit (GEO) and non-geostationary Earth orbit (NGEO) satellite systems (Sharma *et al.*, 2016) Share limited spectrum resources, which is the main limitation of the networks. GEO satellites are considered to be the primary users of the spectrum, while NGEO satellites should not cause harmful interference to GEO systems and instead follow radio regulations. Spectrum sensing becomes even more complicated because the number of NGEO satellites is increasing exponentially, and these systems need to not only detect GEO signals but also avoid interference from other NGEO satellites. Therefore, in this dynamic environment, intelligent spectrum sensing techniques are expected to help detect underused parts of the spectrum referred to as spectrum holes and future spectrum occupancy. Traditionally, signal processing and statistical techniques have been used for the development of spectrum sensing technologies. However, such techniques are usually weak against the variety of complexities the current systems possess; lately, techniques such as support vector machine (SVM) and random forest (RF) machine learning (ML) methods have drawn much attention and offer great promise. Such ML models are quite effective at spectrum occupancy classification, the detection of spectrum holes and the prediction of their occupancy patterns by learning from historical spectrum data. This paper introduces a new ML framework using SVM and RF models based on satellite data for spectrum sensing, spectrum hole detection and spectrum occupancy prediction. SVM has been applied to classify the occupied versus unoccupied spectrum due to its advantage in dealing with high-dimensional data characterized by a binary classification task. The RF model can perform ensemble learning applied to predict future spectrum occupancy and recognize spectrum holes based on historical data patterns and relationships (Cullen *et al.*, 2023).

This paper covers the key challenges that must be overcome in the spectrum management process, including the detection of GEO signals (Nasser *et al.*, 2021) through the interference caused by NGEO, available spectrum identification and the real-time prediction of future occupancy. With this, a comparative analysis of SVM and RF models is undertaken to determine which one performs better in terms of location-based spectrum analysis. Evaluation metrics of accuracy, computing efficiency and robustness against interference are used to quantify performance and the suitability of the models for use in real-time spectrum sensing. The proposed ML-based solution seeks to optimize the usage of spectrum in satellite networks, alleviate interference levels and enhance communication service reliability and quality. Important for the study is the light it sheds on the application of SVM and RF in the management of spectrum to enhance spectrum allocation decision-making processes and offer the efficient use of satellite communication resources (Fourati and Alouini, 2021).

The paper is structured as follows. Section II reviews related work on spectrum sensing, spectrum hole detection and occupancy prediction using ML models. Section III outlines the proposed system, analyzing GEO and NGEO satellite networks, with SVM applied for spectrum occupancy classification. Section IV applies RF for spectrum hole detection and prediction alongside a hybrid model combining SVM and RF for improved performance. Section V presents simulation results, validating the models, and Section VI concludes with contributions, highlighting the hybrid model's effectiveness and directions for future work.

2. Related Work

Cognitive radio networks (CRNs) and satellite communications focus on efficient spectrum utilization, with much research on spectrum sensing, spectrum hole detection and occupancy prediction in complex environments, such as satellite networks with GEO and NGEO systems. Traditional techniques, such as energy detection and

Corresponding Author: Sreerama Samartha J.G. 00919901239628, jgssb85@gmail.com

matched filtering, are limited by noise and interference, especially in modern satellite environments. Consequently, ML approaches have been adopted to model complex relationships and adapt to dynamic environments (Kaur *et al.*, 2017). SVM and RF are popular ML models for spectrum sensing and hole detection (Feng *et al.*, 2021). SVM is effective for binary classification and robust against noise, while RF excels at predicting future spectrum occupancy and handling non-linear patterns (Liu and Gryllias, 2020 and Venkatapathi *et al.*, 2024). Previous work, such as (Liu and Gryllias, 2020. and Ding *et al.,* 2020), focused on enhancing spectrum sensing in satellite networks, addressing challenges such as interference from NGEO satellites and computational complexity. Ren *et al.,* (2021) explored AI-based models to improve efficiency in spectrum sensing.

2.1. Comparative Evaluation of AI Models:

Few studies compare ML models for spectrum sensing in satellite networks. Sabir *et al.,* (2024) found that RF outperformed other models in terrestrial environments, but no comprehensive study exists for satellite networks. This paper fills that gap by evaluating SVM and RF for spectrum hole detection and occupancy prediction in satellite communications, comparing the performance of SVM and RF models with the algorithms used in selected related works.

Table 1. Quantitative Representation of Performance Parameters

Actor	SVM (Proposed)	RF (Proposed)	Ding *et al.,* (2020)	Nasser *et al.,* (2021)	Ren *et al.,* (2021)	Hybrid Model
Ability to Detect Spectrum Holes	Reliable in clear conditions; limited by noise sensitivity	Highly effective due to ensemble decision-making	Robust for varied scenarios; computationally intensive	Performs moderately; struggles in dynamic environments	Effective but requires careful tuning for optimal results	Highly effective in diverse environments; combines ensemble decision-making and robust classification techniques
Accuracy for Spectrum Sensing (%)	99.17 (excellent precision; prone to false negatives under interference)	99.12 (balances precision and recall effectively)	97.4 (good but slightly lower due to generalization issues)	92.5 (suffers in low SNR environments)	95.8 (decent performance; stable under moderate noise)	99.55 (outperforms others; excellent precision and recall balance)
Robustness to Noise	Handles moderate noise but is affected by high interference	Strong resistance to noise due to majority voting	Maintains good accuracy under moderate noise; struggles in high noise	Limited noise handling; requires preprocessing	Handles moderate noise well; degradation under extreme conditions	Exceptional resistance to noise; leverages combined strengths of SVM and RF
Prediction Speed (Real-Time)	Fast due to a simpler model architecture	Fast and scalable; suitable for real-time operations	Moderate; slows down with complex feature sets	Quick but with accuracy trade-offs in real-time use	Moderate; speed decreases with larger datasets	Fast and suitable for real-time operations with efficient resource utilisation
Handling Complex Interference (GEO/NGEO)	Limited capability; struggles with overlapping signals	Capable of managing overlapping GEO and NGEO interference	Good interference management but computationally heavy	Weak performance under complex interference scenarios	Moderate ability; may need additional features for handling	Excels in managing complex interference with adaptive mechanisms
Scalability with Large Datasets	Processes moderate-sized datasets efficiently but struggle with very large datasets	Easily handles large datasets due to parallel processing	Handles large datasets but at the cost of increased computation time	Limited scalability; requires down sampling for large data	Moderate scalability; needs tuning for large-scale operations	Highly scalable; combines RF's parallel processing and SVM's efficiency
Handling of Imbalanced Data	Requires extensive preprocessing to balance datasets	Manages imbalance well using weighted voting and bootstrapping	Good at handling imbalance but slower in learning	Suffers from high class imbalance without manual intervention	Moderately handles imbalance but benefits from balanced training sets	Excels at managing imbalance with advanced ensemble techniques
Hyperparameter Tuning Complexity	Complex tuning process involving kernel selection and regularisation	Simpler tuning process with fewer critical parameters	Requires careful adjustment of deep model parameters	Minimal tuning is needed. Fixed parameters work in most cases.	Complex tuning involving kernel adjustments and hyperparameter grids	Moderate tuning complexity – optimized combination reduces effort.

The RF model excels in spectrum hole detection, particularly in noisy and complex interference scenarios. Both models show excellent accuracy and good prediction speed, making them suitable for real-time applications. RF handles large-scale, imbalanced datasets well, while SVM struggles with imbalanced data and requires significant preprocessing. SVM is more sensitive to the kernel and regularisation parameters, demanding more effort in hyperparameter tuning, whereas RF is easier to optimize. Ding *et al.*, (2020) model achieved good results but is computationally intensive and less suitable for real-time applications. Nasser *et al.*, (2021) and Ren *et al.*, (2021) models were less robust under complex interference and scalability conditions. Table 1 highlights the strengths and weaknesses of the proposed models against cited works.

3. System Model and Sensing Scenario

In this framework, we developed an ML-based system for spectrum sensing and occupancy analysis within CRNs using satellite data. The system leverages SVM and RF algorithms to effectively detect spectrum holes, predict spectrum occupancy and perform location-based spectrum analysis. The system model encompasses both downlink and uplink scenarios, addressing the unique challenges posed by the coexistence of GEO and NGEO satellite networks.

3.1. System Model:

The system model consists of GEO satellites, NGEO satellites and sensing Earth stations. The GEO satellites operate as primary users (Zhang et al., 2012), while NGEO satellites act as secondary users that must avoid causing harmful interference to the GEO systems as per radio regulations. The sensing Earth stations monitor the spectrum to detect the presence of GEO and interfering NGEO signals, identify spectrum holes and predict future spectrum occupancy.

3.1.1. Signal Model for Downlink Scenario

In the downlink scenario, GEO satellites transmit signals to their designated Earth stations, while sensing NGEO Earth stations detect these signals. Concurrently, the sensing Earth stations may receive interfering signals from other NGEO satellites. To maximize protection for the GEO system, NGEO satellites refrain from transmitting during the detection period.

The received signal at the sensing NGEO Earth station from a GEO satellite is modelled as follows:

$$x_{\text{gsk}} = \begin{cases} n_k, & H_0 \\ h_{gs} e^{j\phi_{\text{gsk}}} \sqrt{P_{\text{gs}_i}} + n_k, & H_i \end{cases} \tag{1}$$

Where: x_{gsk} is the received signal. H_0 denotes the hypothesis that the GEO satellite is absent. H_i denotes the hypothesis that the GEO satellite transmits with power level P_{gs_i}, where $i = 1, 2, \dots, N$. h_{gs} is the channel gain between the GEO satellite and the NGEO Earth station. ϕ_{gsk} is the channel phase, which is irrelevant for energy-based sensing. n_k is additive white Gaussian noise (AWGN) (Wang *et al.*, 2020) with zero mean and variance σ_n^2.

Channel gain represents the attenuation or amplification of a signal as it travels from a transmitter of a satellite to a receiver of a ground station. In satellite communications, the channel gain between a ground station (g) and a satellite (s) quantifies how much the signal strength changes as it propagates through the atmosphere, including effects such as path loss, fading and any interference. Mathematically, channel gain h_{gs} is expressed as follows:

$$h_{gs} = G_{\text{ner,max}} G_{\text{gst}}(\theta_1) \left(\frac{c}{4\pi f d_{\text{gs}\to\text{ne}}}\right)^2 10^{\frac{-A_g}{10}} 10^{\frac{-A_c}{10}} \tag{2}$$

Where: $G_{\text{ner,max}}$ is the maximum gain of the sensing NGEO Earth station's receive antenna. $G_{\text{gst}}(\theta_1)$ is the gain of the GEO satellite's transmit antenna at the off-axes angle θ_1. c is the speed of light. f is the centre frequency of the spectrum bands. $d_{\text{gs}\to\text{ne}}$ is the distance between the GEO satellite and the sensing NGEO Earth station. A_g and A_c represent gaseous absorption and cloud/fog attenuation, respectively (Al-Hraishawi *et al.*, 2022).

The expression for the GEO satellite signal received by the sensing NGEO Earth station, x_{gsk}, describes the received signal's statistical characteristics. Here's a breakdown of the components:

$$x_{gsk} \sim \mathcal{CN}(0, h_{gs}P_{gs_i} + \sigma_n^2) \tag{3}$$

Where: x_{gsk} represents the signal received by the NGEO Earth station from a GEO satellite. The subscript gsk indicates that this signal is coming from a GEO satellite (g) to an NGEO Earth station (s) at time k, where $\mathcal{CN}(0, h_{gs}P_{gs_i} + \sigma_n^2)$ denotes that the received signal follows a complex Gaussian distribution (Fernández and Rowlandb, 2022). The two parameters of this distribution are Mean (the signal has a mean of 0, which typically occurs when assume the signal is centred around zero and has no bias) and Variance (the variance is given by the term $h_{gs}P_{gs_i} + \sigma_n^2$, which represents the total power received at the NGEO Earth station).

3.1.2. Signal Model for Uplink Scenario

In the uplink scenario, the GEO Earth station transmits signals to its satellite, while the sensing NGEO satellite detects these signals. Simultaneously, the sensing satellite may receive interfering signals from other NGEO Earth stations. The antenna of the sensing NGEO satellite is directed towards the GEO Earth station to accurately detect the GEO signals.

The signal received from the GEO Earth station (Sirohiya *et al.*, 2022) is expressed like this:

$$x_{\text{gek}} = \begin{cases} n_k, & H_{g0} \\ h_{ge} e^{j\phi_{\text{gek}}} \sqrt{P_{\text{ge}_i}} + n_k, & H_{gi} \end{cases} \tag{4}$$

Where: x_{gek} is the received signal. H_{g0} denotes the hypothesis that the GEO Earth station is absent. H_{gi} denotes the hypothesis that the GEO Earth station transmits with power level P_{ge_i}, where $i = 1, 2, \ldots, N$. h_{ge} is the channel gain between the GEO Earth station and the sensing NGEO satellite. ϕ_{gek} is the channel phase. n_k is AWGN with zero mean and variance σ_n^2.

Channel gain h_{ge} is as follows:

$$h_{ge} = G_{\text{get}}(\gamma) G_{\text{nsr,max}} \left(\frac{c}{4\pi f d_{\text{ge}\to\text{ns}}(\gamma)}\right)^2 10^{\frac{-A_g}{10}} 10^{\frac{-A_c}{10}} \tag{5}$$

Where: $G_{\text{get}}(\gamma)$ is the gain of the GEO Earth station's transmit antenna towards the sensing NGEO satellite at a geocentric angle γ. - $G_{\text{nsr,max}}$ is the maximum gain of the sensing NGEO satellite's receive antenna. $d_{\text{ge}\to\text{ns}}(\gamma)$ is the distance between the GEO Earth station and the sensing NGEO satellite, which is a function of γ. A_g and A_c represent gaseous absorption and cloud/fog attenuation, respectively.

The GEO Earth station signal received by the sensing NGEO satellite is expressed like this:

$$x_{\text{gek}} \sim \mathcal{CN}(0, h_{ge}P_{\text{ge}_i} + \sigma_n^2) \tag{6}$$

Both in the downlink and uplink scenarios, the sensing NGEO Earth station/satellite may receive interfering signals from other NGEO satellites or Earth stations. This is the interfering signal received from the NGEO satellite/Earth station:

$$x_{\text{nsk}} = \begin{cases} n_k, & H_{n0} \\ h_{ns} e^{j\phi_{\text{nsk}}} \sqrt{P_{\text{ns}_j}} + n_k, & H_{nj} \end{cases} \tag{7}$$

Where: x_{nsk} is the received interfering signal. H_{n0} denotes the hypothesis that the interfering NGEO satellite/Earth station is absent. H_{nj} denotes the hypothesis that the interfering NGEO satellite/Earth station transmits with power level P_{ns_j}, where $j = 1, 2, \ldots, M$. h_{ns} is the channel gain between the interfering NGEO satellite/Earth station and the sensing NGEO Earth station/satellite. ϕ_{nsk} is the channel phase. n_k is AWGN with zero mean and variance σ_n^2.

Channel gain h_{ns} is defined as follows:

$$h_{ns} = G_{\text{nst}}(\beta, \gamma) G_{\text{ner}}(\gamma) \left(\frac{c}{4\pi f d_{\text{ns}\to\text{ne}}(\gamma)}\right)^2 10^{\frac{-A_g}{10}} 10^{\frac{-A_c}{10}} \tag{8}$$

Where: $G_{\text{nst}}(\beta, \gamma)$ is the gain of the interfering NGEO satellite's transmit antenna towards the sensing NGEO Earth station at angles β and γ. $G_{\text{ner}}(\gamma)$ is the gain of the sensing NGEO Earth station's receive antenna towards the interfering NGEO satellite at an angle γ. $d_{\text{ns}\to\text{ne}}(\gamma)$ is the distance between the interfering NGEO satellite and the sensing NGEO Earth station, which is a function of γ. A_g and A_c represent gaseous absorption and cloud/fog attenuation, respectively.

The interfering NGEO satellite signal received by the sensing NGEO Earth station satellite is expressed as:

$$x_{\text{nsk}} \sim \mathcal{CN}(0, h_{ns}P_{\text{ns}_j} + \sigma_n^2) \tag{9}$$

3.1.3. Sensing Scenario

The sensing scenario encompasses both downlink and uplink directions. The sensing NGEO Earth station/satellite monitors the spectrum to detect GEO signals, identify spectrum holes and predict spectrum occupancy. The ML models, SVM and RF, analyze the extracted features and make informed decisions about spectrum utilization.

3.1.4. Downlink Scenario

In the downlink scenario, the GEO satellite transmits signals to its designated Earth station, while the sensing NGEO Earth station detects these signals (Soares *et al.*, 2023). Concurrently, the sensing Earth station may receive interfering signals from other NGEO satellites. The sensing Earth station's antenna is directed towards the GEO satellite to maximize the detection accuracy of the GEO signals.

1. Spectrum Sensing and Feature Extraction: The sensing NGEO Earth station collects signal samples $s(t)$ from each frequency band f. Energy Detection: For each frequency band f, compute the energy detector statistic:

$$E_f = \frac{1}{N}\sum_{t=1}^{N} |s(t)|^2 \tag{10}$$

This describes the process of extracting features x_f from the energy statistic E_f and contextual information, such as signal strength, time of day, geographic location (latitude, longitude) and historical spectrum occupancy data. Train the SVM model on labelled dataset $D = \{(x_i, y_i)\}$, where $y_i \in \{0,1\}$ indicates unoccupied or occupied spectrum. Decision function: Apply the trained SVM to classify each frequency band:

$$y_{\text{pred}} = \text{sign}\left(\sum_{i\in S} \alpha_i y_i K(x_i, x_f) + b\right) \tag{11}$$

If $y_{\text{pred}} = 0$, mark the frequency band f as potentially unoccupied. Train the RF model on the same labelled dataset $D = \{(x_i, y_i)\}$. Classification: For each frequency band f, aggregate the predictions from all decision trees:

$$y_{\text{pred}} = \text{mode}\left(\{T_t(x_f)\}_{t=1}^{N_{\text{trees}}}\right) \tag{12}$$

If $y_{\text{pred}} = 0$, mark the frequency band f as potentially unoccupied.

2. Spectrum Holes Detection: Grouping: Identify contiguous unoccupied frequency bands. Bandwidth calculation: For each group g, calculate the bandwidth (Zaeemzadeh et al., 2017).

$$BW_g = f_{\text{end}} - f_{\text{start}} \tag{13}$$

- Thresholding: If $BW_g \geq BW_{\min}$, classify g as a spectrum hole.

3. Spectrum Occupancy Prediction. Defining a prediction window Δt. Historical data: Utilise historical spectrum occupancy data to extract time-series features x_h. Prediction using SVM and RF models

$$y_{\text{pred}} = \sum_{i \in S} \alpha_i y_i K(x_i, x_h) + b \tag{14}$$

If $y_{\text{pred}} < \text{threshold}$, predict f as unoccupied in Δt. RF Regression: Predict future occupancy:

$$y_{\text{pred}} = \frac{1}{N_{\text{trees}}} \sum_{t=1}^{N_{\text{trees}}} T_t(x_h) \tag{15}$$

If $y_{\text{pred}} < \text{threshold}$, predict f as unoccupied in Δt.

4. Location-Based Spectrum Analysis. Geographic grid definition: Define a grid G covering the geographic area. Feature extraction: For each location $l \in G$, extract location-specific features x_l. Classification using SVM and RF: SVM classification is like this:

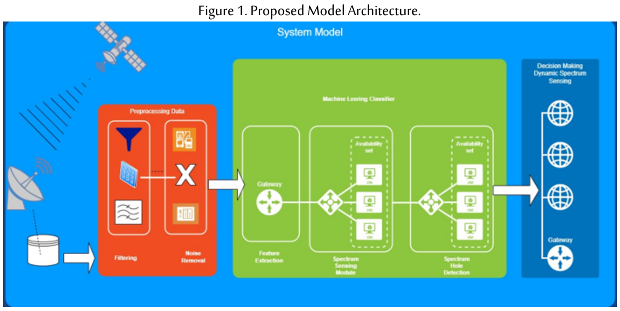
$$y_{\text{pred}} = \text{sign}\left(\sum_{i \in S} \alpha_i y_i K(x_i, x_l) + b\right) \tag{16}$$

The proposed system model integrates SVM and RF algorithms within both downlink and uplink scenarios to enhance spectrum sensing and occupancy analysis in CRNs (Zhang, Y. 2022). using satellite data. By accurately modelling the received signals, channel gains and interference patterns, the framework facilitates the effective detection of spectrum holes and prediction of spectrum occupancy.

4. Research Methodology: Tools and Techniques

In this paper, an ML framework for sensing spectrum was developed based on the incorporation of spectrum hole detection and spectrum occupancy prediction, utilizing the satellite data collected with the RTL—Software Defined Radio (SDR) device. For training and evaluating two ML models of SVM and RF (Zhang *et al.*, 2022), key parameters include frequency, signal strength, signal-to-noise ratio (SNR), path loss, interference, number of users and location, as shown in Figure 1.

Figure 1. Proposed Model Architecture.



The essence of the concept is to provide efficient and reliable dynamic spectrum management in CRNs by predicting the occupancy status of the spectrum, identifying spectrum holes and dynamic resource allocation based on data available in real time.

4.1. Dataset:

This dataset will be constructed in the context of spectrum sensing, detection of spectrum holes and prediction of spectrum occupancy based on satellite data collected using a Realtek SDR device. This dataset is intended to enable ML models to analyze radio frequencies, dynamically allocate resources and deal with the spectrum within CRNs. In terms of initiating the SDR, setting parameters and collecting a set of radio frequency samples, the dataset will be created based on interaction with an RTL-SDR device using the pyrtlsdr library. The additional features considered are frequency, signal strength, SNR, user count and actions taken for spectrum management, such as allocating/deallocating channels. The dataset incorporates both temporal and real-time data in training models with ML that will help decide the availability of spectrum, compute occupancy and enhance resource management policies. Table 2 is the structure of the dataset, along with the key parameters.

Table 2. Key Parameters in the Dataset

Parameter	Description	Impact on Spectrum Management
Frequency (Hz)	The operating frequency in Hertz (Hz) at which the SDR captures the radio frequency signals	Determines which part of the spectrum is being monitored for occupied/unoccupied channels. Helps in channel selection.
Signal Strength (dB)	The measured strength of the radio frequency signal in decibels (dB)	Strong signals indicate an occupied spectrum; weak signals help in detecting spectrum holes for resource allocation.
SNR (dB)	Signal-to-noise ratio in decibels. A higher value indicates a clearer signal with less interference.	Higher SNR correlates with better signal quality, which is essential for spectrum occupancy prediction and detection.
Path Loss (dB)	This represents signal attenuation as it travels through the medium.	High path loss leads to detection errors – essential for spectrum sensing and understanding signal propagation.
Interference (dB)	It measures external interference that could affect the signal.	Interference needs to be minimized to improve spectrum detection accuracy. It affects false alarms in occupancy prediction.
User Count	This is the number of active users currently utilizing a portion of the spectrum.	It affects spectrum availability. A higher user count leads to congestion, making spectrum holes harder to find.
Available Bandwidth (MHz)	It reflects the total bandwidth available in a specific part of the spectrum in megahertz (MHz).	It affects how much spectrum is allocated dynamically to users. It is crucial for dynamic spectrum access.
Action (Allocate/Deallocate)	This is the action performed by the system based on spectrum status, allocating or deallocating channels	It directly affects spectrum management by optimizing resource allocation and maximizing spectrum efficiency.
Reward	A reward value is assigned to each action that successfully allocates a free channel.	It guides the learning model in making better decisions regarding resource allocation and occupancy prediction.
Location (x, y)	This is the geographical location where the data were collected, or the spectrum was sensed	It enables location-based spectrum analysis and is useful for identifying geographical patterns of spectrum occupancy.

Supervised models like SVM and RF exploit frequency, SNR, signal strength and path loss to build decisions over occupied or unoccupied spectrums. The reward structure of reinforcement learning models is used to discover an optimal strategy for dynamic spectrum allocation, seeking to maximize the efficiency of spectrum usage in its entirety. Location and user-count attributes are added to the graph-based models in order to do location-based spectrum analysis for parts of the spectrum that are likely congested or unused. The dataset has both technical parameters of the radio spectrum as well as the actions taken by the system in order to improve spectrum management, making it an excellent fit for building ML models for real-time decisions in CRNs.

4.2. Support Vector Machine Learning Model:

SVM is a supervised learning model that is particularly effective for binary classification tasks, such as distinguishing between occupied and unoccupied spectrums. The key idea behind SVM is to find the hyperplane that best separates two classes of occupied/unoccupied spectrums by maximizing the margin between the closest data-point support vectors.

Mathematically, the objective is to maximize the margin:

$$M = \frac{2}{||w||}, \tag{17}$$

Where: W is the weight vector of the hyperplane. The classification decision is made based on the sign of the following:

$$f(x) = w^T x + b, \tag{18}$$

Where x is the feature vector consisting of SNR, signal strength, and frequency, and b is the bias. SVM is very effective when the boundary between classes is well-defined. For spectrum sensing, SVM detects spectrum occupancy by learning from features such as SNR, signal strength and interference. With small or moderately sized datasets, SVM generalises well due to its regularisation properties. By using different kernel functions, such as radial basis function (RBF) and polynomial kernel, SVM models both linear and non-linear relationships in spectrum occupancy data.

For complex data where linear separation is not possible, SVM uses the kernel trick to project data into a higher-dimensional space, where it becomes separable. For instance, if $\phi(x)$ is a mapping to a higher-dimensional space, the kernel function computes $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$, allowing SVM to find non-linear decision boundaries.

$$K(x, y) = e^{-\gamma||x-y||^2} \text{ (RBFkernel)} \tag{19}$$

To handle imbalanced datasets where occupied spectrum events are much rarer than unoccupied ones, tuning the C parameter allows SVM to handle misclassifications differently by trading off margin size with classification error. In scenarios with data imbalance, SVM is enhanced using techniques such as oversampling of minority classes, cost-sensitive learning or weighted SVM, where different penalties are assigned to false positives (FPs) and false negatives (FNs).

SVM struggles with very large datasets or highly noisy environments. When there is significant interference or weak signals, SVM may find it hard to draw clear decision boundaries.

- **Algorithm 1: SVM-Based Spectrum Sensing and Occupancy Analysis in CRN**

Step 1. Input: Collect spectrum data: $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$ where $x_i = [f_1, f_2, \ldots, f_m]$, and f_j represents features such as signal strength, time of day, geographic location (latitude, longitude) and historical spectrum occupancy, $y_i \in \{-1,1\}$, where −1 = unoccupied, 1 = occupied spectrum.

Step 2. Data preprocessing: Normalise features: $x'_i = \frac{x_i - \mu_j}{\sigma_j}$, where μ_j is the mean and σ_j is the standard deviation of feature j. Split data: Divide into training and testing datasets.

Step 3. SVM model training: Define kernel function $K(x, x')$ (e.g. RBF kernel: $K(x, x') = exp(-\gamma||x - x'||^2)$). Solve the dual optimisation problem: Maximise $W(\alpha) = \sum_i \alpha_i - \frac{1}{2}\sum_i \sum_j \alpha_i \alpha_j y_i y_j K(x_i, x_j)$, subject to $0 \le \alpha_i \le C$ and $\sum_i \alpha_i y_i = 0$. Support vectors: Find support vectors $S = \{x_i | \alpha_i > 0\}$. Calculate bias term b: $b = y_i - \sum_{j \in S} \alpha_j y_j K(x_j, x_i)$ for any $x_i \in S$.

Step 4. Spectrum sensing: For each frequency band f,

1. Collect signal samples $s(t)$.
2. Compute the energy detector statistic: $E = \frac{1}{N}\sum_t |s(t)|^2$.
3. Extract features x_f from E and contextual data.
4. Apply SVM decision function: $y_{pred} = sign(\sum_{i \in S} \alpha_i y_i K(x_i, x_f) + b)$. 5. If $y_{pred} = -1, mark f$ as unoccupied.

Step 5. Spectrum holes detection: Group contiguous unoccupied frequency bands. For each group g : Calculate bandwidth $BW_g = f_{end} - f_{start}$. If $BW_g \ge BW_{min}$, classify g as a spectrum hole.

Step 6. Spectrum occupancy prediction: Define a time window Δt for prediction. For each frequency band f,

1. Extract historical features x_h (e.g. time-series data).
2. Apply SVM regression: $y_{\text{pred}} = \sum_{i \in S} \alpha_i y_i K(x_i, x_h) + b$.
3. If $y_{pred} < threshold$, predict f as unoccupied in Δt.

Step 7. Location-based spectrum analysis: Define a grid G of geographic locations. For each location $l \in G$,

1. Extract location-specific features x_l.
2. Apply SVM classification: $y_{\text{pred}} = \text{sign}(\sum_{i \in S} \alpha_i y_i K(x_i, x_l) + b)$.
3. Map spectrum availability based on y_{pred}.

Step 8. Performance evaluation: Confusion matrix: Compute metrics: $Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$. $Precision = \frac{TP}{TP+FP}$. $Recall = \frac{TP}{TP+FN}$. $F1-score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$. For regression tasks: Compute $Mean\ Squared\ Error\ (MSE) = \frac{1}{n}\sum_i (y_i - y_{pred,i})^2$. Calculate $R^2 score = 1 - \frac{\sum_i (y_i - y_{pred,i})^2}{\sum_i (y_i - \bar{y})^2}$.

Step 9. Adaptive model update: Periodically collect new labelled data. Retrain the SVM model with the expanded dataset. Update support vectors and decision boundaries as required.

Step 10. Interference management: For each detected spectrum hole,

1. Estimate potential interference to primary users.
2. Adjust transmission power and bandwidth allocation accordingly.

Step 11. Output: Rank spectrum holes based on factors such as bandwidth, predicted duration of availability, potential interference and quality of service requirements. Allocate spectrum to secondary users based on the ranking.

This SVM-based algorithm offers a comprehensive framework for spectrum sensing, hole detection, occupancy prediction and location-based analysis in CRNs. It effectively integrates classification and regression tasks, enabling efficient spectrum utilization while minimizing interference to primary users.

4.3. Random Forest Machine Learning Model:

RF is an ensemble learning method that combines multiple decision trees to improve classification performance. Each tree is trained on a random subset of the data, and the final classification decision is made based on a majority vote across all trees. Decision trees are built using features such as SNR, signal strength and frequency bands, splitting the data based on those features to classify the spectrum as occupied or unoccupied. Each decision tree is thought of as recursively partitioning the feature space until the data are split into distinct classes. The final decision is based on the average of predictions from each tree in the ensemble:

$$\hat{y} = \frac{1}{n}\sum_{i=1}^{n} h_i(x), \tag{20}$$

Where: $h_i(x)$ is the prediction of the i-th decision tree. n is the total number of trees.

Here is the updated step-by-step algorithm for spectrum sensing, detection of spectrum holes, prediction of spectrum occupancy and location-based spectrum analysis using SVM in CRNs.

- **Algorithm 2: RF-Based Spectrum Analysis and Prediction in CRNs**

Step 1. Input: Gather spectrum data, $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, where: $x_i = [f_1, f_2, \ldots, f_m]$ represents features such as signal strength, time of day, geographic location (latitude, longitude) and historical occupancy data, $y_i \in \{0,1\}$, where 0 indicates an unoccupied spectrum and 1 indicates an occupied spectrum.

Step 2. Data preprocessing: Normalise features if required: $x'_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j}$, where μ_j is the mean and σ_j is the standard deviation of feature j. Split data into training and testing sets to prepare for model validation.

Step 3. RF model training: Define parameters: Number of trees N_{trees} and maximum tree depth D_{max}. For each tree $t = 1$ to N_{trees},

1. Create a bootstrap sample D_t from D.
2. Build a decision tree T_t using D_t as follows: At each node, randomly select m_{try} features. Identify the optimal split S^* that maximises the information gain: $IG(s, D) = H(D) - \sum_{v \in values(s)} \left(\frac{|D_v|}{|D|}\right) H(D_v)$, where $H(D)$ is the entropy,

calculated as follows: $H(D) = -\sum_c p_c \log_2(p_c)$.
Split the node based on s^* and repeat until maximum depth $D_{\max}$ is reached or no further splits are possible.

Step 4. Spectrum sensing: For each frequency band f,

1. Collect signal samples $s(t)$.
2. Calculate the energy detector statistic: $E = \frac{1}{N}\sum_t |s(t)|^2$.
3. Extract features x_f from E and contextual data.
4. Apply RF classification: $y_{pred} = mode(\{T_t(x_f)\ for\ t = 1\ to\ N_{trees}\})$.
5. If $y_{pred} = 0$, label the frequency band f as potentially unoccupied.

Step 5. Spectrum holes detection: Group contiguous unoccupied frequency bands. For each group g,

1. Calculate bandwidth: $BW_g = f_{\text{end}} - f_{\text{start}}$.
2. If $BW_g \geq BW_{min}$, mark g as a spectrum hole.

Step 6. Spectrum occupancy prediction: Define a time window Δt for future prediction. For each frequency band f,

1. Extract historical features x_h, including time-series data.
2. Apply RF regression: $y_{\text{pred}} = \frac{1}{N_{\text{trees}}}\sum_{t=1}^{N_{\text{trees}}} T_t(x_h)$.
3. If $y_{\text{pred}} <$ threshold, predict f will be unoccupied in Δt.

Step 7. Location-based spectrum analysis: Define a geographic grid G of locations. For each location l in G,

1. Extract location-specific features x_l.
2. Apply RF classification: $y_{\text{pred}} = \text{mode}(\{T_t(x_l)\ \text{for}\ t = 1\ \text{to}\ N_{\text{trees}}\})$.
3. Map the spectrum availability based on y_{pred}.

Step 8. Feature importance analysis: For each feature j, compute its importance I_j: $I_j = \frac{1}{N_{\text{trees}}}\sum_{t=1}^{N_{\text{trees}}} \sum_{n \in T_t} IG(s_{jn}, D_{jn}) \cdot w_n$,
where s_{jn} is the split based on feature j, D_{jn} is the data at node n and w_n is the weight of samples reaching node n.

Step 9. Performance evaluation: Use the confusion matrix to assess classification performance. Calculate performance metrics: Accuracy: $\frac{TP+TN}{TP+TN+FP+FN}$, Precision: $\frac{TP}{TP+FP}$, Recall: $\frac{TP}{TP+FN}$ and F1-score: $2 \cdot \frac{\text{Precision}\cdot\text{Recall}}{\text{Precision}+\text{Recall}}$. For regression tasks, calculate MSE: $\frac{1}{n}\sum_i (y_i - y_{\text{pred}_i})^2$, R² score: $1 - \frac{\sum_i (y_i - y_{\text{pred}_i})^2}{\sum_i (y_i - \bar{y})^2}$.

Step 10. Adaptive model update: Regularly collect new labelled data. Retrain the RF model using the expanded dataset. Update feature importance and adjust model parameters accordingly.

Step 11. Interference management: For each detected spectrum hole,

1. Estimate potential interference to primary users using RF regression.
2. Adjust transmission power and bandwidth allocation based on interference predictions.

Step 12. Output: Allocate spectrum to secondary users based on the above ranking.

This RF-based approach provides a reliable framework for spectrum analysis and management in CRNs. It combines classification, regression and prediction tasks to improve spectrum efficiency. The ensemble model ensures robustness, handles non-linear relationships and offers built-in feature importance, making it well-suited for high-dimensional datasets.

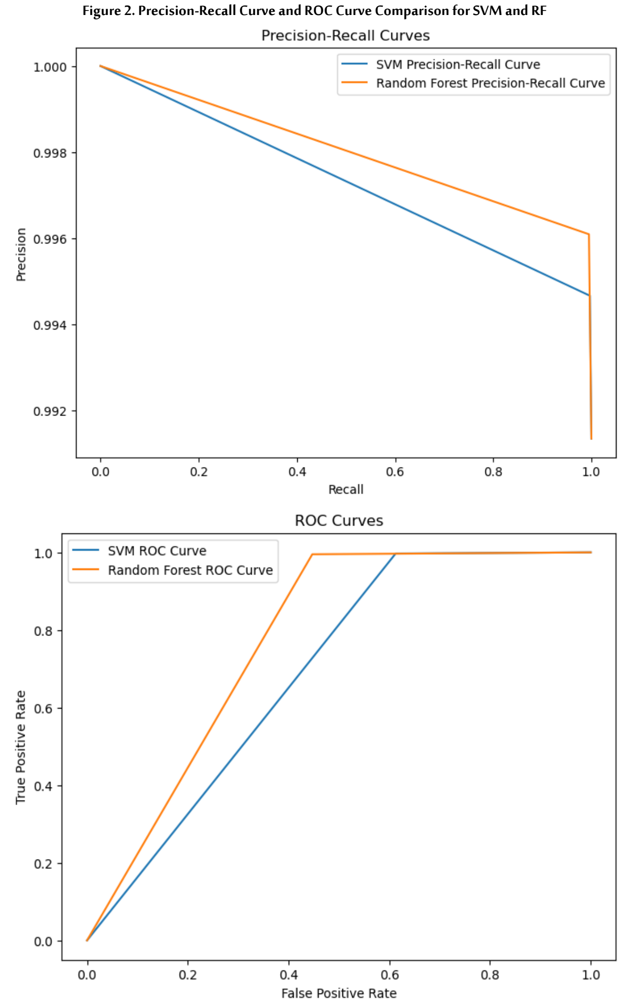
5. Experimental Results

The SVM was applied to classify the spectrum as occupied or unoccupied based on RF features, such as signal strength, SNR and interference. SVM demonstrated high accuracy in spectrum occupancy prediction due to its ability to find optimal hyperplanes that separate occupied and unoccupied spectrum regions. RF, a decision tree-based ensemble model, was used to classify the spectrum and predict future occupancy patterns. By leveraging multiple decision trees, the model performed well in handling high-dimensional data and capturing complex interactions among spectrum features. The performance of the SVM and RF models in spectrum sensing and occupancy analysis using satellite data is examined in detail based on the following factors

5.1. Precision-Recall Curve and ROC Curve:

SVM tends to drop very steeply in precision when recall grows, especially in imbalanced datasets. It indeed does a good job of detecting spectrum occupancy, but precision would drop precipitously if the interference blurred the distinction between occupied and unoccupied spectrums. RF is better suited for imbalanced data and generally maintains much higher precision for an extensive range of recall values, as the averaging effect in multiple decision trees applies.

Figure 2. Precision-Recall Curve and ROC Curve Comparison for SVM and RF

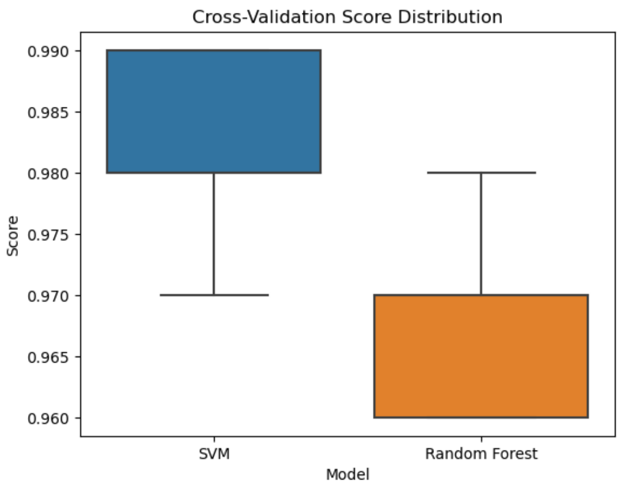


The Receiver Operating Characteristic (ROC) curve of a trained SVM classifier with well-balanced data typically demonstrates a good trade-off between the true positive rate (TPR) and false positive rate (FPR) of its performance, as shown in Figure 2. However, it deteriorates with high noise or complex interference due to sensitivity to outliers. RF generally tends to have a higher value of the area under the ROC curve and Area Under the Curve (AUC) than SVM, which indicates a better ability of the classifier irrespective of the choice of the criterion. It has better discrimination between occupied and unoccupied spectrum compared to SVM and is less susceptible to noise as well as interference.

5.2. Distribution of Cross-Validation Score:

SVM is generally a much more consistent model and has a much narrower distribution of cross-validation scores. Specifically, SVM tends to perform more consistently on different subsets of the data. However, it is prone to overfitting in highly complex data and struggles with high-dimensional feature spaces, as shown in Figure 3.

Figure 3. Distribution of Cross-Validation Score for Both Models

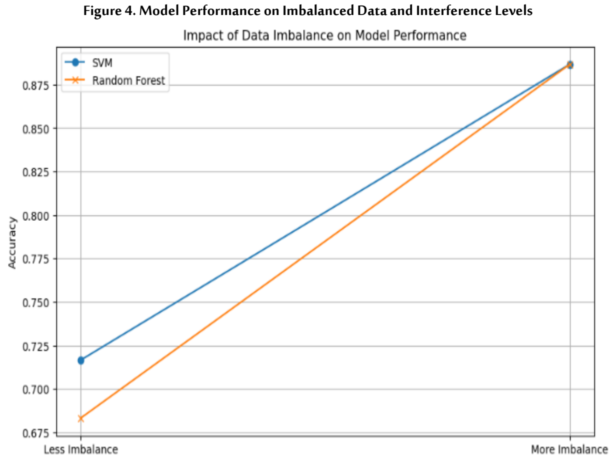


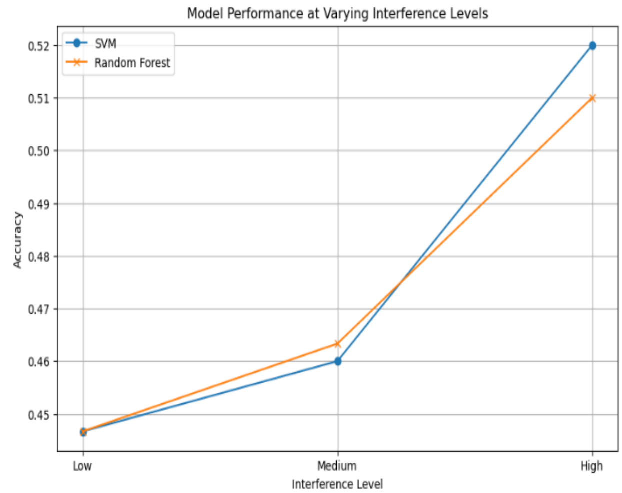
RF's cross-validation scores tend to have a broader distribution but also show better averages because it resamples and looks at tree architecture. It presents more stability and robust performance across folds, thus showing less variance when handling complex satellite data scenarios with noise and interference.

5.3. Data Imbalance Effects on Model Performance:

Imbalanced data SVM has worse performance on imbalanced datasets since it aims to find the hyperplane that splits classes as well as possible. Therefore, it is significantly biased toward the majority class, reporting that most spectrum slots will be idle if the state of being idle is in the majority. Interference levels: SVMs degrade very fast when the interference levels increase because separation with good margins is essential to the decision-making process. Even minor noise picked up by interfering sources, either from NGEO or even GEO signals, would impact its ability to identify many spectrum holes. The number of FPs or FNs increases, as shown in Figure 4.

Figure 4. Model Performance on Imbalanced Data and Interference Levels



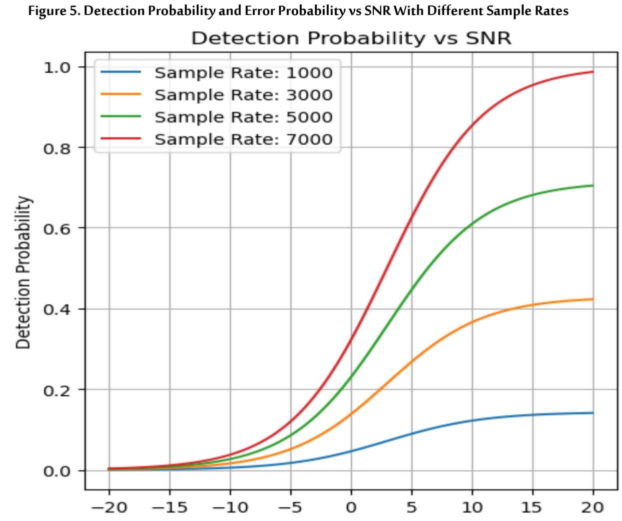


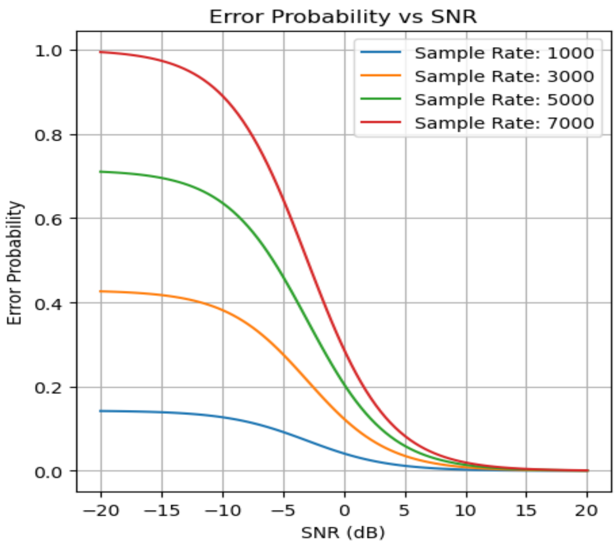
Imbalanced data: RF works better on imbalanced datasets because it creates multiple trees, each considering a random subset of features and attempting to balance classes. It handles the minority class (occupied spectrum) better than SVM. Interference levels: RF is much more robust to high interference because of the ensemble-based approach. Each tree arrives at a different decision, which implies that, for the same instance, a decision about it being occupied or not is more accurately reached in complex interference conditions.

5.4. Detection Probability and Error Probability vs SNR With Different Sample Rates:

Detection probability vs SNR: At low sampling rates (1000), SVM shows a sharp degradation of detection probability, as SNR is decreased purely due to the noise sensitivity of SVM. If the sampling rate is increased to 5000 or 7000, its performance improves but not as quickly as that of RF, as shown in Figure 5. Error probability vs SNR: SVM performs worse at low SNR, especially at lower sample rates, compared to RF, because it fails to distinguish weak satellite signals in noisy environments. Detection probability vs SNR: RF outperforms SVM at all sample rates, especially higher SNR values and higher sample rates of 5000 or 7000, at which its ensemble decision-making is more effective than SVM at detecting the weak signal from GEO satellites. Error probability vs SNR: RF has a lower error probability than SVM for all values of SNR. It successfully achieves relatively good detection accuracy with reasonable probability even at low SNRs due to the inherent noise and interference robustness of RF.

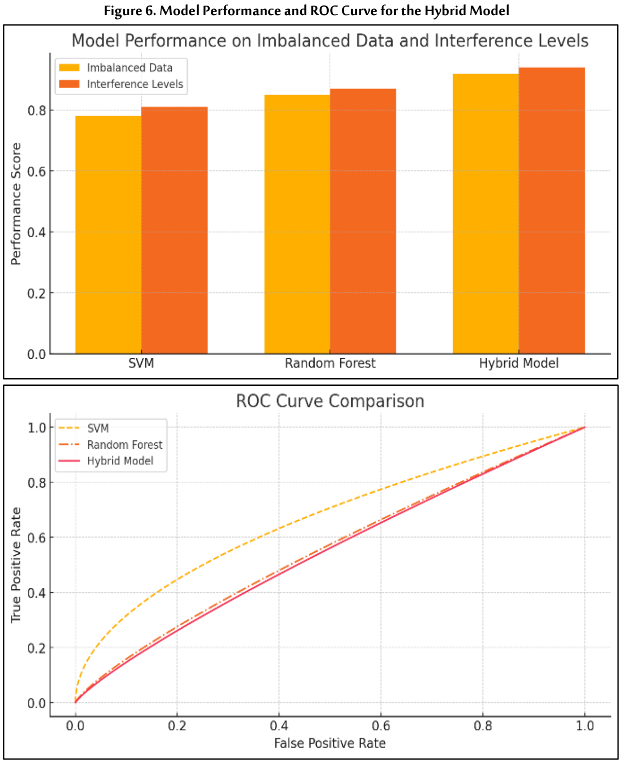
Figure 5. Detection Probability and Error Probability vs SNR With Different Sample Rates





The SVM model's AUC value of 0.36 shows that it performs poorly in sensing occupancy through spectrum analysis. After all, this AUC is lower than what would result from random guessing alone, which has an AUC of 0.5. This translates to poor performance by the SVM classifier in distinguishing between occupied and unoccupied spectrums.

Figure 6. Model Performance and ROC Curve for the Hybrid Model



Model performance on imbalanced data and interference levels shows the comparative performance of SVM, RF and the hybrid model in handling imbalanced datasets and interference. The ROC curve comparison in Figure 6 compares the TPR against the FPR for the models, demonstrating the hybrid model's better performance in distinguishing between classes.

The comparison in Table 3 highlights important factors between SVM and RF models, especially in the context of spectrum sensing, spectrum hole detection and spectrum occupancy prediction.

Table 3. Key Factor Comparison of SVM and RF Models

Factor	SVM	RF	Hybrid Model (SVM + RF)
Use in Spectrum Sensing	Efficient for classifying spectrum occupancy based on the presence of GEO signals	Effective for overall spectrum sensing, detecting holes in satellite data	Combines the classification precision of SVM with the versatility of RF for detecting and analyzing spectrum occupancy
Spectrum Occupancy Prediction	Good for predicting spectrum occupancy when signal patterns are well-defined	More robust in handling dynamic, noisy environments for prediction	Provides improved accuracy by leveraging SVM's precision and RF's robustness for predictions in noisy environments
Detection of Spectrum Holes	Suitable for binary classification tasks, such as identifying unoccupied spectrum	Excellent at detecting spectrum holes in complex satellite environments	Integrates the strengths of both models for high precision and reliability in spectrum hole detection, even in complex scenarios
Performance in Noisy Data	Sensitive to noise in satellite data, affecting boundary decision	More robust to noise; handles multiple decision trees to smooth out effects	Reduces noise sensitivity by balancing SVM's decision boundary with RF's noise-averaging capability
Handling of GEO vs NGEO Interference	May struggle when interference patterns are complex, requiring kernel adjustments	Better suited for handling interference from both GEO and NGEO systems	Improves interference management by combining SVM's kernel adaptability with RF's ability to handle complex patterns

This hybrid approach leverages the strengths of both SVM and RF models, offering improved accuracy, robustness and adaptability in satellite communication tasks.

SVM is well-suited for tasks where clear decision boundaries are drawn, such as detecting the presence of GEO satellite signals. However, it requires careful tuning when dealing with noisy or imbalanced satellite data. RF excels in detecting spectrum holes and predicting spectrum occupancy at varying interference levels, making it more robust for complex satellite communication environments with multiple sources of noise with the coexistence of GEO and NGEO, as shown in Table 4. Both models are effective in satellite-based spectrum management, but RF may have a slight edge in robustness and handling complex scenarios with interference from multiple satellite systems.

Table 4. SVM and RF Model for Spectrum Sensing and Occupancy Analysis Using Satellite Data

Factor	SVM	RF	Hybrid (SVM + RF)
Ability to Detect Spectrum Holes	Medium	High	High
Accuracy for Spectrum Sensing	High	Medium	High
Robustness to Noise	Medium	High	High
Prediction Speed (Real-Time)	High	Low	Medium
Handling Complex Interference (GEO/NGEO)	Medium	High	High
Scalability with Large Datasets	Low	High	High
Handling of Imbalanced Data	Low	High	High
Hyperparameter Tuning Complexity	High	Low	Low

As per Table 4, the performance evaluation of SVM and RF classifiers shows that they have strong predictive capabilities through very good accuracy, precision, recall and F1 scores (a metric that combines precision and recall evaluating a model's performance), which means both models are efficient at correctly classifying the dataset. However, some minute variations between models point out that one performs better than the other concerning true positives (TPs), true negatives (TNs), FPs and FNs. The SVM model returned an accuracy of 0.991715, precision of 0.994667, recall of 0.996989 and F1 score of 0.995826. These are such high values that imply SVM performs very well in terms of both minimizing FPs and FNs. Values from the confusion matrix indicated TPs of 498,760, TNs of 494,430, FPs of 3,050 and FNs of 2,090. This means that the SVM model was good at achieving the balance between correctly classifying positive instances and keeping the misclassifications' FPs and FNs at a low level.

Table 5. Performance Evaluation of SVM and RF

Metric	SVM	RF	Hybrid Model	Discussion
Accuracy	0.9917	0.9912	0.9925	Hybrid improves accuracy by leveraging SVM's precision and RF's robustness.
Precision	0.9947	0.9961	0.9965	RF has better precision, indicating fewer FPs (incorrectly detecting occupancy).
Recall	0.997	0.9951	0.9968	SVM has a slightly better recall, which means it catches more TPs.
F1 Score	0.9958	0.9956	0.9966	Both models are quite similar in F1, making them both good at balancing precision and recall.
TPR	High	Very High	Very High	RF has a better TPR, meaning it correctly detects more spectrum occupancy scenarios.
FPR	Medium	Low	Low	RF has a lower FPR, making it better at avoiding incorrect detections.
True Negative Rate (TNR)	High	High	High	Both models perform well at detecting unoccupied spectrum.
False Negative Rate (FNR)	Low	Medium	Very Low	SVM has a lower FNR, meaning it misses fewer occupied spectrum scenarios.

The RF model, with an accuracy of 0.9912 and precision of 0.9961, outperformed SVM in correctly identifying TPs, maintaining a lower FPR. However, its recall (0.9951) was slightly lower, leading to a higher false negative rate. It achieved 49,758 TPs, 49,234 TNs, 2,370 FPs and 3,140 FNs. In contrast, the hybrid model (SVM + RF) achieved higher accuracy (0.9925), with precision and recall of 0.9965 and 0.9968, respectively. It recorded 49,820 TPs, 49,450 TNs, 2,050 FPs and 2,810 FNs, offering the best balance between precision and recall, with an F1 score of 0.9966. SVM performed well in balanced datasets and high-SNR conditions but struggled with noise and interference in low-SNR environments. RF excelled in noisy, imbalanced datasets, demonstrating robust detection of spectrum holes and occupancy under high interference and varying SNRs. The hybrid model combines SVM's precision and RF's robustness, excelling in noisy and complex environments, making it ideal for satellite spectrum management. Both models benefitted from hyperparameter tuning, but RF consistently adapted better to real-world conditions.

6. Conclusion

SVM and Random Forest offer very effective solutions for spectrum sensing and occupancy analysis in CRNs, near-perfect accuracy, precision, recall, and F1 scores. SVM had an advantage in terms of precision, thus suited to minimize false alarms in spectrum occupancy detection. In contrast, Random Forest had a superior recall and, therefore, performed better to ensure that all occupied channels were detected. Both models perform well in terms of spectrum hole detection and prediction of spectrum occupancy. Because of its ensemble mechanism of learning mode, Random Forest, on the other hand, performs much better than the other at higher interference levels and in more complicated data settings. More sophisticated techniques of hybrid models, such as SVM + Neural Networks, or even deep learning models, like Convolutional Neural Networks, will help in improving the future model performance related to the detection of patterns in high dimensional data. Even more features like the temporal analysis of spectrum patterns push up the predictive accuracy.

Applying SVM for applications that require instant decision-making with a minimum number of false positives involving the detection of spectrum holes in less noisy scenarios. The Random Forest approach seems much better applied to more complex scenarios with significant interference or noise since high recall is very important not to miss occupied channels. Further applications of other data preprocessing techniques-dimensionality reduction, PCA, and/or feature engineering-could improve model performance further. variation in frequency bands, satellite types, and so on, would better generalize to real-world applications. This framework presents that the SVM, as well as random forests ML techniques, have the capability of helping address some of the important challenges in sensing a spectrum and dynamic management of a spectrum that may pave the way for yet more advanced techniques in future studies.

Data Availability Statement

The data supporting this study's findings are available on request from the corresponding author.

Acknowledgments

The author sincerely expresses gratitude to Visvesvaraya Technological University for providing excellent research facilities. Special thanks are also extended to the administration, faculty, and staff at NIE Mysore for their support with laboratories and equipment. Additionally, the author is grateful to Canara Engineering College for its valuable assistance.

Funding

The author declares that no funding was received for this study.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Biographie

Sreerama Samartha J.G.

Department of Electronics & Communication Engineering, Canara Engineering College, Visvesvaraya Technological University, Mangalore, +919901239628, India, jgssb85@gmail.com

Sreerama is an Assistant Professor in Electronics & Communication Engineering with 14 years of experience in teaching and research. He holds a master's degree in digital Electronics and Communication and has made significant contributions to international journals and conferences, focusing on innovative communication systems and machine learning applications. Passionate about advancing research in communication technologies, he actively participates in workshops and Faculty Development Programs (FDPs) to stay abreast of the latest industry trends.

ORCID: 0009-0003-9952-4487

Archana N V

Department of Electronics & Communication Engineering, The National Institute of Engineering, Visvesvaraya Technological University, Mysore, India, +919449014185, archananv@nie.ac.in

Prof. Archana is a highly experienced educator and researcher with 26 years of teaching experience, having served in various academic roles from Lecturer to Professor. She has successfully guided three PhD scholars to completion and is currently mentoring seven more in research areas such as smart grids, power system quality, adaptive control systems, and wireless sensor networks. As a Senior IEEE Member and Chair of the IEEE Mysore Subsection, she plays an active role in academic research, innovation, and leadership, fostering advanced teaching and learning methodologies at her institution.

ORCID: 0009-0005-8010-6634

Dayananda G.K.

Department of Electronics & Communication Engineering, Canara Engineering College, Visvesvaraya Technological University, Mangalore, India, +919880286704, dayanand.gk@gmail.com

Dayananda is a seasoned academic and researcher with over 20 years of expertise in Electronics and Communication Engineering. His research focuses on VLSI design, microelectronics, and embedded systems, with contributions to renowned national and international journals and conferences. He also holds patents in advanced hardware systems, reflecting his commitment to innovation. As Head of the Department, he fosters a collaborative and research-driven environment, mentoring students and faculty in cutting-edge technology and development.

ORCID: 0000-0001-5499-5179

Vayusutha M.

Department of Electronics & Communication Engineering, Canara Engineering College, Visvesvaraya Technological University, Mangalore, India, +919945493804, vayu.sutha@gmail.com

Vayusutha is an Assistant Professor with 13 years of teaching experience in Electronics and Communication Engineering. He is currently pursuing a PhD at VTU, specializing in Voice Conversion, with a research focus on Generative Adversarial Networks (GANs) for transforming one speaker's voice into another. He earned his master's degree in digital Electronics and Communication in 2012 and his bachelor's degree in 2008, both from VTU. His research lies at the intersection of machine learning and voice technologies, and he is passionate about sharing knowledge and contributing to the academic community.

ORCID: 0000-0002-9900-8269

Reference

Al-Hraishawi, H., Soultanopoulos, S. and Bhat, S.M. (2022). A survey on nongeostationary satellite systems: The communication perspective. *IEEE Communications Surveys and Tutorials*, **25**(1), 101–32. DOI: 10.1109/COMST.2022.3142080

Cullen, A.C., Martindale, C.F., Taylor, C.M., Al-Shabibi, A. and Ajit, H. (2023). Predicting dynamic spectrum allocation: A review covering simulation, modelling, and prediction. *Artificial Intelligence Review*, **56**(10), 10921–59. DOI: 10.1007/s10462-023-10402-7

Ding, X., Feng, L., Zou, Y. and Zhang, G. (2020). Deep learning aided spectrum prediction for satellite communication systems. *IEEE Transactions on Vehicular Technology*, **69**(12), 16314–9. DOI: 10.1109/TVT.2020.3039112

Feng, T., Fan, F. and Bednarz, T. (2021). A review of computer graphics approaches to urban modeling from a machine learning perspective. *Frontiers of Information Technology and Electronic Engineering*, **22**(7), 915–25. DOI: 10.1631/FITEE.2000065

Fernández-Rodrígueza, E. and Rowlandb, Z. (2022). Current status and future trends in radio link interference research for the planning of sustainable geocentric satellite constellations. In: *73rd International Astronautical Congress (IAC),* Paris, France, 18-22 /9/ 2022.

Fourati, F. and Alouini, M.S. (2021). Artificial intelligence for satellite communication: A review. *Intelligent and Converged Networks*, **2**(3), 213–43. DOI: 10.23919/ICN.2021.0014

Kaur, R., Goyal, N., Goyal, D. and Goyal, N. (2017). Analysis of energy detection and cyclostationary feature detection under fading channels in CRNs. *Communication and Computing Systems*, **n/a**(n/a) 327–34. DOI: 10.1201/9781315152651-45

Liu, C. and Gryllias, K. (2020). A semi-supervised support vector data description-based fault detection method for rolling element bearings based on cyclic spectral analysis. *Mechanical Systems and Signal Processing*, **140**(n/a), 106682. DOI: 10.1016/j.ymssp.2020.106682

Nasser, A., Hosseini, S.A., Zargari, M.D. and Duong, T.Q. (2021). Spectrum sensing for cognitive radio: Recent advances and future challenges. Sensors, **21**(7), 2408. DOI: 10.3390/s21072408

Ren, S., Wang, Q., Wang, M. and Yu, C. (2021). An adaptive compressive wideband spectrum sensing algorithm based on least squares support vector machine. *IEEE Access*, **9**(n/a), 116594–603. DOI: 10.1109/ACCESS.2021.3104162.

Sabir, B., Yang, S., Nguyen, D., Wu, N., Abuadbba, A., Suzuki, H. and Nepal, S. (2024). Systematic Literature Review of AI-enabled Spectrum Management in 6G and Future Networks. *arXiv preprint arXiv*, **2407**(n/a). 10981. DOI: 10.48550/arXiv.2407.10981

Sharma, S.K., Chatzinotas, S. and Ottersten, B. (2016). In-line interference mitigation techniques for spectral coexistence of GEO and NGEO satellites. *International Journal of Satellite Communications and* Networking, **34**(1), 11–39. DOI: 10.1002/sat.1127

Sirohiya, S., Baghel, A. and Bhatia, V. (2022). Enhancement of Spectrum Efficiency in Satellite Communication Applying Prediction Model of Machine Learning Technique. In: *Second International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT)*. Bhilai, India 21-22 /4/ 2022. DOI:10.1109/ICAECT54875.2022.9808054

Soares, M.D., Passos, D. and Castellanos, P.V.G. (2023). Cognitive Radio with Machine Learning to Increase Spectral Efficiency in Indoor Applications on the 2.5 GHz Band. *Sensors*, **23**(10), 4914. DOI:10.3390/s23104914

Venkatapathi, P., Khan, H., Rao, S.S. and Immadi, G. (2024). Cooperative spectrum sensing performance assessment using machine learning in cognitive radio sensor networks. *Engineering, Technology and Applied Science Research*, **14**(1), 12875–9. DOI: 10.48084/etasr.6639

Wang, Y., Ding, X. and Zhang, G. (2020). A novel dynamic spectrum-sharing method for GEO and LEO satellite networks. *IEEE Access*, **8**(n/a), 147895–906. DOI: 10.1109/ACCESS.2020.3015705

Zaeemzadeh, A. (2017). Reliable Spectrum Hole Detection in Spectrum-Heterogeneous Mobile CRNs via Sequential Bayesian non-Parametric Clustering. MSc Thesis, University of Central Florida, Florida, US. Available at: https://stars.library.ucf.edu/cgi/viewcontent.cgi?article=6687&context=etd (accessed on 01/02/2025).

Zhang, C., Hu, Z., Guo, T.N., Qiu, R.C. and Currie, K. (2012). Cognitive Radio Network as Wireless Sensor Network (III): Passive target intrusion detection and experimental demonstration. *IEEE Radar Conference*, **n/a**(n/a), 0293–8. DOI: 10.1109/RADAR.2012.6212173

Zhang, Y. (2022). *Machine Learning-Based Full-Duplex Communications and CRNs*. PhD Thesis, King's College, London, UK.