Machine Learning based Spectrum Sensing and Spectrum Management using Blockchain

Sarvesh Uttarwar
Department of ETE
Rashtreeya Vidhyalaya College of Engineering
Bengaluru,India
sarveshu.et20@rvce.edu.in

Swastik Ghosh
Department of ETE
Rashtreeya Vidhyalaya College of Engineering
Bengaluru,India
swastikghosh.et20@rvce.edu.in

Aditya Prakash
Department of ETE
Rashtreeya Vidhyalaya College of Engineering
Bengaluru,India
adityaprakash.et20@rvce.edu.in

CH Mani Pujith
Department of ETE
Rashtreeya Vidhyalaya College of Engineering
Bengaluru,India
poornachandramp.et20@rvce.edu.in

Dr. Nagamani K
Department of ETE, Professor
Rashtreeya Vidhyalaya College of Engineering
Bengaluru,India
nagamanik@rvce.edu.in

Abstract— The rapid advancement of wireless technologies like IoT and 5G has exposed the limitations of traditional spectrum management methods, resulting in inefficiencies such as spectrum underutilization and security vulnerabilities. Machine learning (ML)-based spectrum sensing addresses these issues by analyzing and predicting usage patterns, enhancing utilization, and reducing underutilization. Additionally, the centralized nature of traditional systems makes them prone to security breaches and lacks transparency. Blockchain technology mitigates these concerns by offering a decentralized, tamper-proof ledger for managing spectrum allocations. Integrating ML and blockchain optimizes spectrum utilization, enhances security, and maintains the integrity of spectrum management.Countries like the United States, Japan, Korea, China, and Germany have advanced spectral measurement technology using SNR values. The U.S. leads in cognitive radio network research, Japan focuses on urban user identification, South Korea on 5G spectrum allocation, China on reliable sensing amid noise, and Germany on machine learning for spectral accuracy. The model integrates frequency-based spectrum sensing with blockchain for adaptable spectrum allocation. Machine learning enhances utilization and signal identification, while blockchain ensures secure and transparent spectrum management. The Data is preprocessed QoS data for 5G networks and apply ML models to predict vacant bands, evaluating their performance with R-squared scores and mean squared errors.

Keywords— Spectrum management,Machine learning (ML),Blockchain technology,Spectrum sensing,Cognitive radio networks,Signal-to-noise ratio (SNR)

1. INTRODUCTION

The evolution of 5G networks has necessitated the development of advanced data analysis and management systems to ensure optimal performance and resource allocation. This research aims to address two critical aspects of 5G network management: data-driven quality of service analysis and secure spectrum allocation. By leveraging machine learning techniques and blockchain technology, this research provides a comprehensive solution for analyzing service quality and managing spectrum resources efficiently.

Machine learning-based spectrum sensing is transforming RF spectrum management in wireless communication systems. Algorithms like k-nearest neighbors (KNN), random forest, and others significantly improve the accuracy and efficiency of identifying vacant spectrum bands. KNN discerns patterns in RF signals based on similarity to existing data points, while random forest utilizes multiple decision trees to classify signals robustly amidst noise and variability. These methods empower spectrum sensing systems to detect available spectrum bands reliably, optimizing spectrum utilization without compromising communication reliability.

Integrating blockchain technology into spectrum management holds promise for revolutionizing efficiency and transparency in wireless communication systems. Blockchain's decentralized and immutable ledger capabilities can streamline spectrum allocation, usage tracking, and interference management. By utilizing smart contracts, blockchain can automate spectrum leasing agreements, ensuring secure and transparent transactions between stakeholders without intermediaries. This approach enhances spectrum utilization efficiency by reducing administrative overhead and minimizing the risk of spectrum misuse or unauthorized access. Furthermore, blockchain's tamper-proof nature and consensus mechanisms enable real-time verification of spectrum ownership and usage rights, fostering trust among network participants and regulatory bodies alike. Implementing blockchain in spectrum management not only promises operational efficiencies but also lays the groundwork for more equitable and sustainable spectrum allocation practices in increasingly congested RF environments. Section 1 highlights the need to use machine learning based spectrum sensing to find vacant bands, and use of blockchain in spetrum management.

Section 2 reviews literature on machine learning based spectrum sensing, blockchain and dynamic spectrum sensing . Section 3 details the methodology of integrating machine learning-based spectrum sensing with blockchain to provide secured access and ensure integrity also highlights the intricate features of the dataset. Section 4 presents results of both the models. Section 5 concludes that combining machine learning based spectrum sensing with blockchain offers a promising solution.

Key Terms and Libraries Used:

Pandas is essential for data manipulation, offering efficient structures like DataFrames for large datasets and supporting tasks such as cleaning and statistical analysis. NumPy provides foundational support for scientific computing with its powerful ndarray for array operations and mathematical functions. Matplotlib creates high-quality visualizations with extensive customization, while Seaborn, built on Matplotlib, simplifies statistical visualizations. TensorFlow, a platform developed by Google, excels in machine learning and deep learning, providing high-level APIs and efficient computation across multiple hardware types.

Spectrum management Involves spectrum allocation, licensing, and enforcement of regulations to ensure fair and efficient use of radio frequencies across different users and services, balancing conflicting demands while minimizing interference.

Machine learning (ML) Employs algorithms and statistical models to enable computers to learn from and make predictions or decisions based on data, without being explicitly programmed. ML finds applications in spectrum prediction, cognitive radio systems, and optimizing spectrum usage.

Blockchain technology Utilizes decentralized and distributed digital ledgers to record transactions securely and transparently across multiple computers. In spectrum management, blockchain can enhance trust and efficiency in spectrum auctions, leasing, and ensuring compliance with spectrum regulations.

Spectrum sensing Refers to the capability of radio devices to detect and monitor RF signals in their vicinity. Spectrum sensing techniques include energy detection, cyclostationary feature detection, and matched filtering, crucial for dynamic spectrum access and avoiding interference.

Cognitive radio networks Networks where devices intelligently perceive their operational environment, adapt their parameters, and behavior in real-time to optimize spectrum utilization. Cognitive radios enable efficient spectrum sharing and can dynamically adjust to varying RF conditions.

2. LITERATURE REVIEW

The paper [1] explores Dynamic Spectrum Access (DSA) and network connectivity challenges in wireless communication, proposing blockchain technology for enhanced management. Blockchain-enabled DSA and network slicing offer benefits in spectrum management, access control, and connectivity. Platforms like Ethereum and Hyperledger Fabric are evaluated for smart contract suitability, demonstrating blockchain's potential in B5G and 6G networks.

The paper [2] investigates deep neural networks (DNN) for spectrum sensing in 6G networks, highlighting MLPs for classification, CNNs for spatial analysis, LSTM networks for temporal dependencies, and combined CNN-LSTM architectures for comprehensive feature extraction. Autoencoders aid in anomaly detection, enhancing spectrum sensing accuracy and adaptability.

The paper [3] examines smart radio technology for spectrum utilization in IoT, proposing a blockchain-based dynamic spectrum access framework to mitigate vulnerabilities of traditional central fusion center architectures. Integrating blockchain with machine learning (ML) in Cognitive Radio IoT (CR-IoT) networks enhances spectrum management's security and efficiency.

The paper [4] proposes a blockchain-based method for detecting malicious users in cognitive radio networks (CRN), addressing security challenges from the rapid growth of IoT devices. Blockchain ensures the integrity of spectrum management, achieving 100% accuracy in detecting malicious users and enhancing the reliability of CRNs.

The paper [5] suggests integrating neural network-based techniques with spectral detection algorithms to improve spectrum sensing in low SNR conditions. Using CNN, ResNet, and LSTM networks, this approach enhances detection performance and strengthens spectrum sensing reliability in dynamic wireless scenarios.

The paper [6] highlights the limitations of traditional static spectrum allocation models and advocates for dynamic spectrum management to meet modern communication demands. Blockchain technology is proposed to enhance spectrum efficiency, security, and reliability through a decentralized management framework.

The paper [7] addresses the need for efficient wireless communication solutions for IoE devices and small-scale technologies in 5G and B5G networks. It explores Smart Spectrum Sharing (SS) leveraging AI, reviewing SS approaches, and analyzing advancements like radio recognition and blockchain to improve spectrum utilization and network performance.

The paper [8] compares deep learning-based approaches for spectrum sensing in cognitive radios, using SP features and NN architectures like MLP, CNN, ResNet, and LSTM. Experiments show MLP and ResNet offer the best detection performance, especially when trained with combined SP features.Methodology

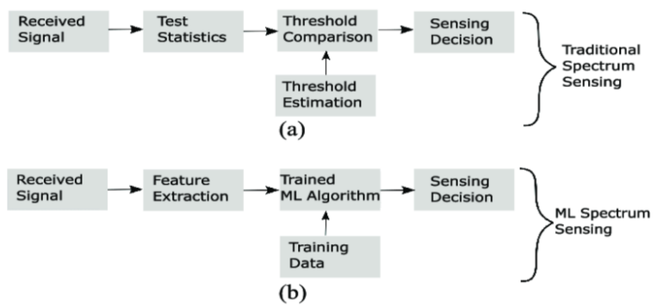


Figure 1 *depicts two approaches to spectrum sensing: traditional & ML-based*

Traditional Spectrum Sensing

In traditional spectrum sensing (Figure 1a), the system receives a radio signal and compares it to a predefined threshold. If the signal strength exceeds this threshold, it is classified as a signal; otherwise, it is classified as noise. This straightforward method relies on predefined criteria to determine the presence or absence of a signal.

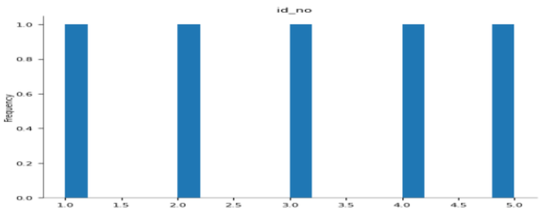
Machine Learning-Based Spectrum Sensing

In machine learning-based spectrum sensing (Figure 1b), the system receives a radio signal and extracts relevant features such as signal strength, power spectral density, and statistical properties. These features are used to train a machine learning model with labeled data, allowing the model to learn patterns indicative of a signal. Once trained, the model analyzes unknown signals and makes decisions based on the learned patterns, enhancing detection accuracy and adaptability compared to traditional methods.

	Timestamp	id_no	Application_Type	Signal Stren	Latency	Res_Allo	person	Req1_band(Mhz)	Alloc1_band(Mhz)	low1_freq	up1_freq	Vacant1
0	09-03-2023 10:00	1	Video_Call	-75	30	70%	John	2.50	3.750	0.250	0.75	1.250
1	09-03-2023 10:00	2	Voice_Call	-80	20	80%	Peter	0.25	0.300	0.025	0.25	0.050
2	09-03-2023 10:00	3	Streaming	-85	40	75%	Rahul	1.25	1.500	0.750	6.25	0.250
3	09-03-2023 10:00	4	Emergency_Service	-70	10	90%	John	0.25	0.375	0.250	0.50	0.125

Figure 2 *first five rows of dataset after cleaning it*

The dataset consists of timestamped records detailing various communication sessions within a network, each identified by a unique id_no and characterized by parameters such as Application_Type, Signal_Strength (in dB), Latency (in ms), Res_Allo (resource allocation efficiency), person (user), Req1_band(Mhz) and Alloc1_band(Mhz) (requested and allocated bandwidth), low1_freq and up1_freq (frequency range bounds), and Vacant1 (spectrum availability).



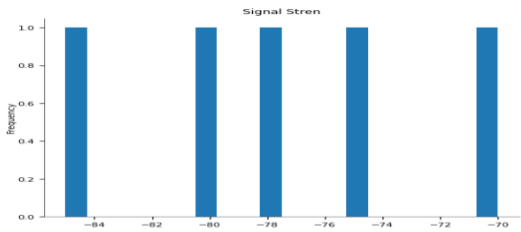
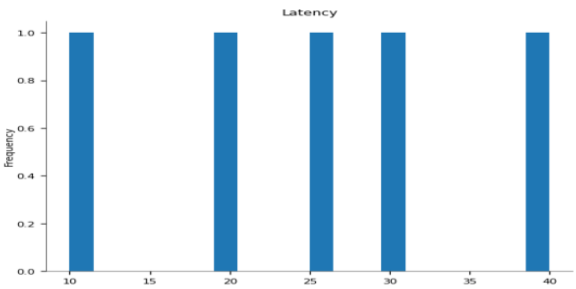


Figure 3 *shows the features of dataset*



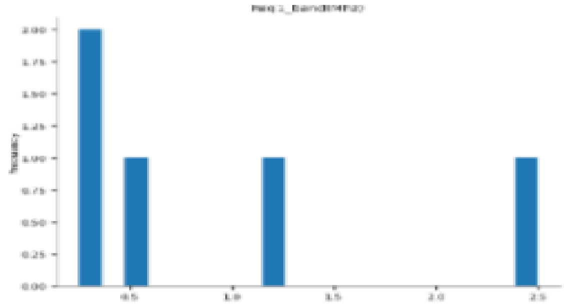


Figure 4 *shows the features of dataset*

The Timestamp column provides temporal context, while Application_Type categorizes network activities like Video Calls, Voice Calls, Streaming, and Emergency Services. Signal_Strength and Latency offer insights into network quality and user experience. Res_Allo measures resource utilization efficiency, and Vacant1 indicates spectrum availability. Together, these parameters enable a thorough analysis of network performance, resource allocation, and optimization strategies. The fig 3 and fig 4 shows the data set after preprocessing the data with frequency on y- axis and id_no. ,signal stren , latency , Req1_band in respective x-axis.

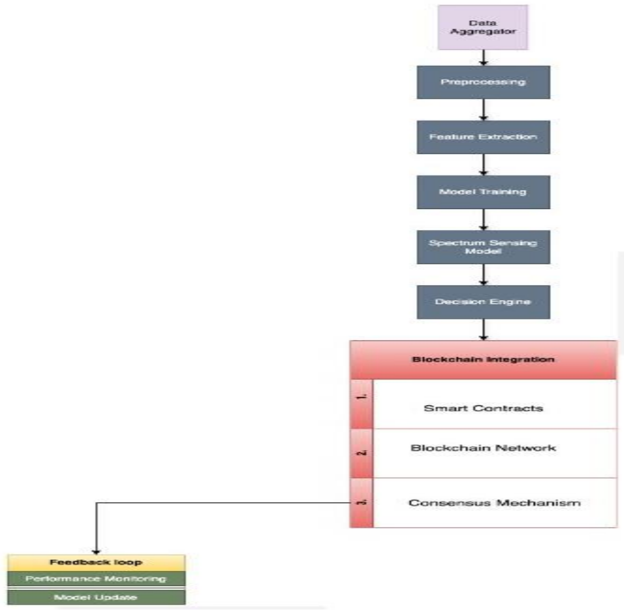


Figure 5 *Methodology*

The figure 5 shows the methodology which has been implemented to deploy the model.

The work commences by accessing and preprocessing data sourced from Google Drive to optimize its usability. This preprocessing involves the removal of redundant columns such as 'Req_Band(in Mbps)', 'Alloc_Band(in Mbps)', 'low_freq', 'up_freq', 'Vacant', and 'Timestamp'. Additionally, percentage values in columns like 'Signal Stren', 'Latency', and 'Res_Allo' are converted from string to float format for better manipulation and analysis. Subsequently, the dataset undergoes exploratory data analysis (EDA), where descriptive statistics and visualization techniques, including bar charts, are employed to discern the dataset's characteristics and distribution of application types, providing valuable insights for subsequent analysis.

After data exploration, a filtering mechanism is introduced to isolate pertinent records within specified frequency ranges, based on user-input lower and upper frequency bounds. Feature engineering techniques come into play, involving the creation of dummy variables for categorical features such as 'Application_Type' and label encoding for 'person' columns, facilitating effective representation of the data for subsequent machine learning analysis. Furthermore, numerical feature standardization ensures uniform scaling, thereby simplifying model training across various regression algorithms, including Linear Regression, K-Nearest Neighbors (KNN) Regression, Decision Tree Regression, and Random Forest Regression. The dataset is then partitioned into training and testing sets (80:20) to evaluate model performance, with predictions made for the target variable 'Vacant1' based on the test data.

The visualization of results encompasses a diverse array of graphical representations. Bar charts are utilized to illustrate average vacant bandwidth by application type and person, providing a clear visual depiction of allocation trends. Line charts are employed to demonstrate frequency trends, shedding light on frequency utilization patterns across different applications. Scatter plots visualize the relationship between actual and predicted values for each model, offering insights into model accuracy and performance. Additionally, histograms are utilized to display the distribution of vacant bandwidth, facilitating a deeper understanding of bandwidth availability.

Models are updated with adjusted datasets to incorporate new data and insights gained from ongoing analysis, thereby ensuring their adaptability to changing radio environments. These adjustments may involve retraining models with updated datasets or fine-tuning model parameters to enhance accuracy and performance. Spectrum allocation is a critical aspect of efficiently managing wireless communication networks. To address the complexities associated with this task, a blockchain-based system has been developed. This system harnesses the inherent benefits of blockchain technology, including transparency, security, and decentralization, to facilitate the effective allocation and monitoring of spectrum usage within a network.

The system employs a Spectrum Class to represent individual frequency bands within the network. This class encapsulates attributes such as ID, lower frequency bounds, and upper frequency bounds, providing a comprehensive representation of each spectrum segment. Additionally, spectrum allocations are represented using the SpectrumAllocation Class, which contains essential details regarding the allocation of a frequency band to a specific user, including spectrum information, the identity of the owner, and the allocated start and end times.

Central to the blockchain system is the Block Class, which serves as the foundational building block encapsulating spectrum allocations. This class incorporates attributes such as hash, nonce, previousHash, allocations, and timestamp. The SpectrumBlockchain Class orchestrates the management of the blockchain, overseeing critical functionalities such as block creation, mining, allocation validation, and conflict resolution. Initialization of the blockchain involves the creation of the genesis block, marking the inception of the blockchain without any spectrum allocations.

New spectrum allocations are seamlessly incorporated into the blockchain by adding pending allocations to the system. Mining processes convert these pending allocations into a new block, maintaining the blockchain's integrity and continuity. Procedures for retrieving allocations from the blockchain and verifying the chain's integrity are essential for managing the system. Furthermore, the system carefully validates new allocations to prevent conflicts with existing ones, ensuring there is no overlapping spectrum usage or time conflicts.

To ensure fair and efficient spectrum utilization, the system employs a conflict resolution mechanism. This mechanism automatically rejects new allocations if they overlap with existing ones, promoting equitable and non-overlapping spectrum usage among users. An example scenario demonstrates the practical functionality of the blockchain system in managing spectrum allocations. This example showcases the creation of spectrum objects, their allocation to users, the validation of allocations, and their subsequent integration into the blockchain.

3. Result

	Emergency_Service	File_Download	IoT_Temperature	Online_Gaming	Streaming	Video_Call	Video_Streaming	VoIP_Call	Voice_Call	Web_Browsing
0	False	False	False	False	False	True	False	False	False	False
1	False	False	False	False	False	False	False	False	True	False
2	False	False	False	False	True	False	False	False	False	False
3	True	False	False	False	False	False	False	False	False	False
4	False	False	False	True	False	False	False	False	False	False

***Figure 6** is a screenshot of a table containing data*

The Figure 6 describes the table containing data , where each application type is converted from string to binary data . In the context of the work, each application type undergoes a transformation process wherein it is converted from string format to binary data. This conversion is integral to the data preprocessing stage and serves to enhance the efficiency and effectiveness of subsequent analyses and model training procedures. By encoding categorical variables such as application types into binary representations, the system can better interpret and utilize these features within machine learning algorithms.The conversion process involves generating binary variables corresponding to each unique application type present in the dataset. For instance, if the dataset contains application types like "Video_Call," "Emergency_Service," and "Online_Gaming," each of these categories is assigned a binary value. This binary conversion effectively changes categorical data into a format that can be understood and processed by machine learning models.

Enter the lower frequency: 0.25
Enter the upper frequency: 0.75

Application_Type	person	Actual	Linear Regression	KNN Regression	Decision Tree Regression	Random Forest Regression
Video_Call	John	0.325	0.355544	0.335	0.325	0.325000
Emergency_Service	Karan	0.025	0.079805	0.035	0.025	0.025750
Video_Call	Ignatius	0.325	0.353034	0.325	0.325	0.325000
Online_Gaming	Pibarel	0.075	0.064732	0.075	0.075	0.075000
Video_Call	Tanay	0.325	0.345247	0.325	0.325	0.325000
Video_Call	Ignatius	0.300	0.466935	0.615	0.300	0.367750
Video_Call	Rahul	0.325	0.343843	0.325	0.325	0.325000
Online_Gaming	Peter	0.075	0.060550	0.075	0.075	0.075000
Emergency_Service	Ignatius	0.025	0.033441	0.035	0.025	0.033125
Video_Call	Rahul	0.325	0.346359	0.325	0.325	0.325000
Emergency_Service	Karan	0.000	-0.003405	0.000	0.000	0.000000
Online_Gaming	Vikas	0.075	0.122548	0.075	0.075	0.079375
Online_Gaming	Karan	0.075	0.091797	0.075	0.075	0.075000
Video_Call	Karan	0.500	0.399946	0.405	0.500	0.526500

***Figure 7** comparison of various machine learning algorithms.*

The user interface prompts users to input lower and upper frequency values, guiding them through the process of specifying the frequency range of interest, likely in Megahertz (MHz). Within the dataset, the "Application_Type" column identifies the type of application requesting resources, with "Emergency_Service" highlighted as one such application. The accompanying "person" column lists individuals assigned to emergency services, including names like John, Alex, Rahul, and others. The primary target variable to be predicted by the regression models is outlined, representing the actual value of interest. Additionally, the predictions generated by each regression model are presented for each data point, offering insights into the performance of these models in predicting the target variable. Each row in the dataset corresponds to a specific combination of application type and individual, providing a comprehensive overview of the predicted values across different scenarios and regression techniques.

In the analysis of various application types, observations reveal that Random Forest and Decision Tree Regression consistently perform well, often providing exact matches to actual values. While Linear Regression occasionally overestimates or underestimates, KNN Regression shows reasonable performance. Overall, Random Forest and Decision Tree models emerge as reliable choices for regression tasks in this context, demonstrating strong predictive capabilities across different application scenarios.

Figure 8 and 9 illustrates scatter plots comparing predicted versus actual vacant bandwidth for two regression models: Decision Tree Regression and Random Forest Regression. Each point represents a prediction, with the x-axis indicating actual bandwidth and the y-axis indicating predicted bandwidth. Ideally, points should align on the 45-degree line (y = x), indicating perfect predictions.

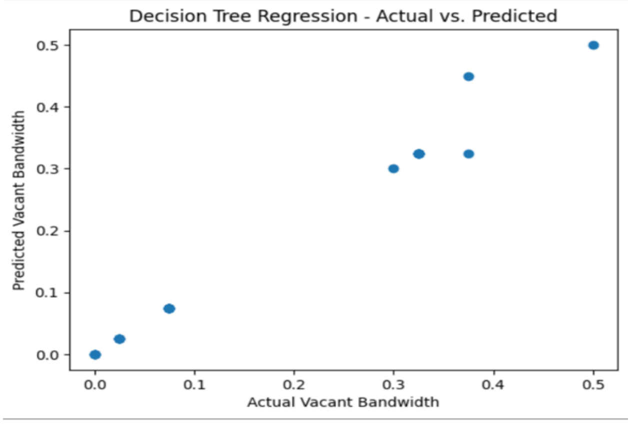


Figure 8 *Scatter Plot in Decision Tree*

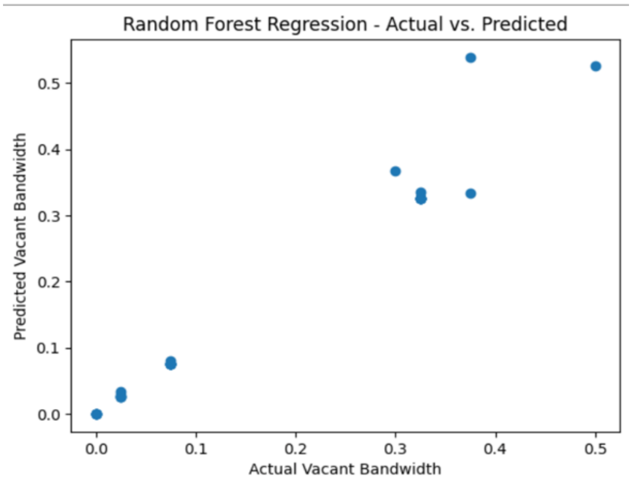


Figure 9 *Scatter plot in Random Forest*

In the Decision Tree Regression plot, points roughly align with this line but deviate, especially at higher bandwidths, suggesting some inaccuracies, particularly with increased bandwidth. Conversely, the Random Forest Regression plot shows tighter clustering around the 45-degree line, indicating higher accuracy and consistency. This suggests that Random Forest Regression, by averaging predictions from multiple decision trees, better captures data patterns, reducing variance and enhancing overall prediction accuracy.

Figure 10 *Blockchain results*

Figure 10 displays a command line window capturing messages related to blockchain transactions. The text indicates the system is verifying allocation validity, followed by adding valid allocations to the blockchain. Each transaction generates a unique block hash, created using a cryptographic hash function to ensure data integrity. Blockchain technology, as described, is a secure and distributed ledger system resistant to tampering. This interaction with the blockchain network demonstrates the process of securely adding valid transactions to the ledger.

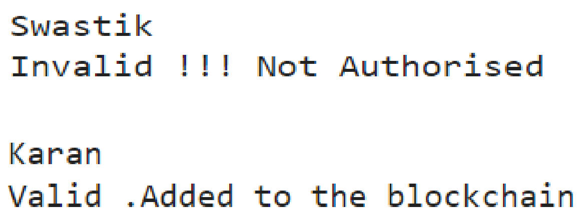


Figure 11 *Blockchain results*

The Figure 11 shows "Allocation is invalid, Not Authorised." indicating that the blockchain system detected an invalid allocation and is attempting to remove it from the blockchain with only the authorized users are allocated with unutilized spectrum band.

4. Conclusion

The research presents a comprehensive analysis and implementation for managing 5G Quality of Service (QoS) data and validating spectrum allocations using blockchain technology. It begins with data preprocessing, visualization, and filtering based on frequency ranges. Various regression models are trained to predict vacant bandwidth, with Random Forest Regression achieving the highest accuracy at 62.06%. Blockchain technology is then employed to ensure no overlap and maintain integrity in spectrum allocations, with valid allocations mined and added to the blockchain. Key insights include the importance of preprocessing, the varied insights from different machine learning models, and the efficacy of blockchain in ensuring valid, conflict-free spectrum management. User interaction ensures allocation validity against the blockchain, providing a robust and secure mechanism for 5G QoS management.

5. Future Scope

The future scope of the machine learning-based approach for spectrum management involves refining feature selection techniques, optimizing model performance through hyperparameter tuning, and exploring ensemble learning methods for more accurate predictions. Additionally, deep learning architectures offer potential for uncovering complex usage patterns. Real-time spectrum allocation protocols can be developed to efficiently manage dynamic usage, while incentive structures can encourage user participation and data sharing within the blockchain-based system. Enhanced blockchain security measures will ensure the integrity of spectrum allocation transactions. Integration of these enhancements promises a more efficient and secure spectrum management system, combining machine learning and blockchain technology to optimize resource allocation and maintain transparency.

REFERENCES

[1] T. A. Hall, R. Caromi, M. Souryal, and A. Wunderlich, ‘‘Reference datasets for training and evaluating RF signal detection and classification models,’’ in Proc. IEEE Globecom Workshops (GC Wkshps), Waikoloa, HI, USA, Dec. 2019, pp. 1–5.

[2] S. N. Syed, P. I. Lazaridis, F. A. Khan, Q. Z. Ahmed, M. Hafeez, V. Holmes, I. P. Chochliouros, and Z. D. Zaharis, ‘‘Deep learning approaches for spectrum sensing in cognitive radio networks,’’ in Proc. 25th Int. Symp. Wireless Pers. Multimedia Commun. (WPMC), Oct. 2022, pp. 480–485.

[3] V. Krishnakumar, P. Savarinathan, P. Savarinathan, T. Karuppasamy, and A. Jayapalan, "Machine learning based spectrum sensing," in *Proc. 2022 International Conference on Computer Communication and Informatics (ICCCI)*, Jan. 2022, pp. x-x. doi: 10.1109/ICCCI54379.2022.9740824.

[4] F. A. Khan, S. N. Syed, and P. I. Lazaridis, "Deep Neural Networks for Spectrum Sensing," IEEE Transactions on Neural Networks, vol. 2023..Pg no: 58-81

[5] [4] M. S. Miah, M. S. Hossain, and A. G. Armada, "Machine Learning-based Malicious Users Detection in Blockchain-Enabled CR-IoT Network for Secured Spectrum Access," IEEE Internet of Things Journal,, 2023..Pg no: 81-115

[6] M. K. Luka, O. U. Okereke, E. E. Omizegba, and E. C. Anene, "Blockchains for Spectrum Management in Wireless Networks: A Survey," in IEEE Communications Surveys & Tutorials , Pg no: 120-132

[7] S. S. Chandra, A. Upadhye, P. Saravanan, S. Gurugopinath and R. Muralishankar, "Deep Neural Network Architectures for Spectrum Sensing Using Signal Processing Features," 2021 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER), Nitte, India, 2021, pp. 129-134

[8] Carlos Serôdio & José Cunha & Guillermo Candela & Santiago Rodriguez & Xosé Ramón Sousa & Frederico Branco, 2023. "The 6G Ecosystem as Support for IoE and Private Networks: Vision, Requirements, and Challenges," Future Internet, MDPI, vol. 15(11), pages 1-32