RESEARCH ARTICLE WWW.IJCCTS.ORG

Protecting Distributed Ledgers from Advanced Persistent Threats Using SVM-Based Blockchain Security

Kiruba Buri R[1]*, K. Swaminathan[2]

[1]Department of CSE, University College of Engineering , Pattukottai, Tamil Nadu

[2]Department of ECE, SRM Institute of Science and Technology, Ramapuram Campus, Chennai, Tamil Nadu

Keywords:
Blockchain,
Distributed ledger, APT,
Machine Learning,
Cyber security,
Threat identification,
Threat patterns,
Decentralized architecture,
Real time protection,
Predictive analytics Modeling

Author's Email id:
srikirubaburi@gmail.com,
swaminathan.vinoth@gmail.com

Author's Orcid id:
0000-0002-0293-5809,
0000-0002-8116-057X

DOI: 10.31838/IJCCTS.13.02.02

Received : 13.02.25

Revised : 20.03.25

Accepted : 18.04.25

Abstract

This work focuses on multi-dimensional approach to incorporate the Support Vector Machine (SVM) models with Blockchain to secure distributed ledger against APTs. The classifying and high pattern recognition ability of SVM makes the proposed framework easily capture and neutralize malicious activities in the blockchain networks in realtime. The distribution of the blockchain technology and use of machine learning for predictive modeling guarantees a hard-coded countermeasure against new forms of cyber threats. As such, this work is centered on how these technologies can be integrated in harmony: attempting to enhance the accuracy of threat identification without compromising the functionality of the blockchain. This implementation shows the possibility of achieving strong, secure and scalable applications in different applications domains, and so make a way forward for upcoming decentralized cybersecurity solutions.

How to cite this article: Kiruba Buri R, Swaminathan K (2025). Protecting Distributed Ledgers from Advanced Persistent Threats Using SVM-Based Blockchain Security. International Journal of communication and computer Technologies, Vol. 13, No. 2, 2025, 11-17

Introduction

Machine learning application integrated with the blockchain is considered a suitable approach to solving cybersecurity problems in the distributed network. Blockchain as a technology that relies on distributed processing and record-keeping and which has shown to be particularly successful in ensuring that data remains and transactions are secure.. However, its susceptibility to new habits APTs, changed cyber threats require powerful solutions. Artificial neural networks, most preferably Support Vector Machines (SVM) are renowned to be effective in pattern and anomaly detection on large data volumes and thus should be considered for improving blockchain safety.

Much research has been conducted on the complementarity of the mentioned technologies.

For example, researchers suggested frameworks based on machine learning models to detect, prevent malicious actions inside the blockchain networks and maintain the reliability and integrity of distributed ledgers.[1] Other research focuses on the contributions of SVM in fraudulent transactions' identification to improving consensus algorithms in blockchain systems.[2] Machine learning for optimizing the application of blockchain on the basis of the scalability and reliability has also been discussed in the latest work .[3] Additionally, some of the studies have investigated about the merging supervisor learning enhances intrusion identification and mitigates unauthorized threats to decentralized networks infrastructure.[4]

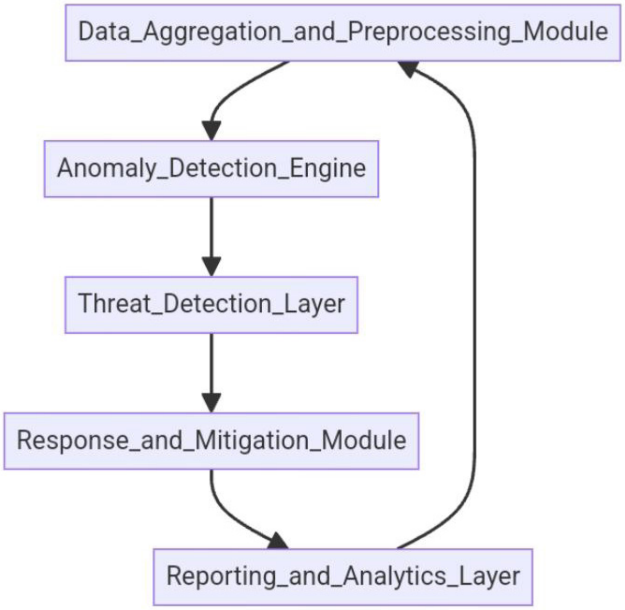
Data_Aggregation_and_Preprocessing_Module
Anomaly_Detection_Engine
Threat_Detection_Layer
Response_and_Mitigation_Module
Reporting_and_Analytics_Layer

Fig. 1: General flow of cybersecurity framework.

The cybersecurity framework which is used to fight APTs uses a layered model to counter the threats successfully: as can be seen in the following figure 1. In its essence, the Blockchain-Based Security Framework safeguards transactions by providing the reliability of the recorded information. The framework also encompasses the Threat Detection Layer the leverage of enhanced machine learning algorithms to analyze network activities with a view of identifying anomalous activities and or suspicious behaviors. Newly generated or already existing data from Iot-devices or network traffic are accumulated by the Data Aggregation and Preprocessing Module and preprocessed providing the Anomaly Detection Engine with information about deviations from the standard behavior that can be associated with threats. If one or more of these conditions is triggered, the Response and Mitigation Module implements countermeasures and sends alerts to the system administrator, negates the threat automatically, and performs a blockchain analysis to identify the source of the attack. The fact is that the Reporting and Analytics Layer offers recommendations for improving the security measures and maintaining them as up to date as possible. This integrated system provides secure end to end protection system against all sorts of cyber threats.

Furthermore, integration of blockchain decentralized approach with machine learning predictors has been found effective in addressing cyber risks in distributed systems.[5] The authors have also observed the use of real-time anomaly detection to mitigate complex cyber threats in blockchain systems using machine learning.[6] That has been made possible by the flexibility of SVM in detecting structures in blockchain networks to give early signals of security risks.[7]

In addition, research studies have proposed the method for secure communication in the blockchain-based IoT system employing the use of machine learning techniques confirming the cross-domain applicability of IoT.[8] Other works focusing on the combination of blockchain technology with machine learning for threat intelligence have shown how the precision of threat identification and neutralization can be enhanced.[9] Finally, the progressive emergence of more efficient machine learning algorithms has provided the basis for the creation of new safe and effective forms of blockchain, as well as their reliable protection against contemporary cyber threats.[10]

Literature Survey

The combination of blockchain and machine learning has attracted the concern of the community to improving cybersecurity measures for mitigating sophisticated threats. Several articles have discussed the possibility of applying blockchain in cases where the results are guaranteed, and security incidents are unified without the involvement of third parties. For example, one work examined the blockchain-based intrusion detection system with the aid of machine learning for anomalies detection with emphasis on enhancing the threat identification in real-time situations.[11] Another study was conducted to explain the usage of blockchain for protection of IoT communication link by exploring machine learning technology for scalability of distributed networks.[12]

Also, the integration of blockchain with artificial intelligence has been used with positive results that put an end to external and unauthorized infiltration.[13]

A remarkable exploratory study explored the application of neural network to improve consensus algorithms of blockchain, which elaborated how it could boost security while maintaining effectiveness.[14] In addition, authors have deployed Support Vector Machines (SVM) in blockchain network to perform the task of fraudulent transaction detection and ensuring the correctness of the ledger.[15] Other researches have also analysed the possibility of using machine learning models in smart contract security and ensuring that the validation processes are automatic and cannot be changed.[16] Another important work suggests to design a blockchain secure communication protocol with machine learning to identify new types of cyber threats.[17] Likewise, the contributions of reinforcement learning for enhancing the blockchain network resilience to advanced persistent threat were investigated comprehensively in related research.[18] Studies carried out on the application of federated learning model with a blockchain technology proved that it was effective in maintaining data security and also promoting efficient threat identification.[19] Lastly, the implementation of machine learning combined with blockchain solutions was described in a study as the way to get predictive analysis for taking proper security measures.[20]

Proposed Framework

The presented work aims to develop a novel system that combines blockchain and machine learning for improving cyber security of distributed networks from APTs. The execution starts with data acquisition where data is obtained from IoT devices, network traffic logs and blockchains. Data preprocessing comes next it involves data cleansing and normalization of data for use in machine learning. Thereafter, feature extraction is carried out to obtain pertinent attributes such as traffic habits of the network and consumer actions. The system uses machine learning algorithms for identifying anomalies which can indicate threat. The real-time monitoring of the network allows constant identification of behavior trends deviation. When an attack is identified, the response and mitigation phase triggers other countermeasures to be initiated including system termination or IP address elimination. It contains blockchain audit for all actions that take place so that it is more secure and transparent. A notification system is established that alarms administrators of a possible risk, encouraging a swift reaction. The system provides reports of the observed cycles and the steps taken to prevent the identified problems from occurring, which contributes to adapting security approaches. Last, feedback provides substantial reinforcement of learning and enhancement that reflects detection and response actions. This flow guarantees strong, more importantly, admissible, and dynamic protection as depicted in the figure 2.

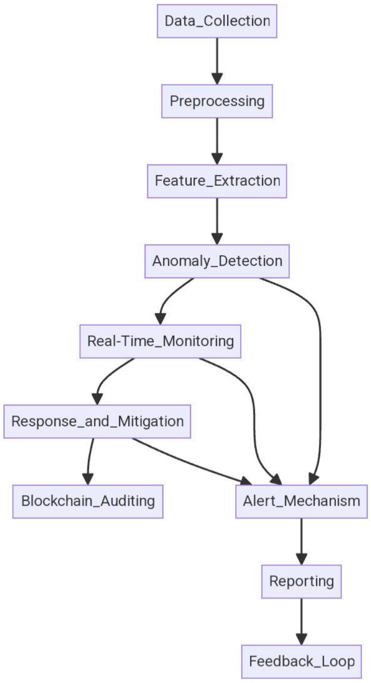

Fig. 2: Proposed framework flow mechanism

The introduced system integrates the blockchain approach along with state-of-the-art machine learning with the aim to increase cybersecurity and protect against APTs in distributed networks. The system is a linear, several-stage program incorporating data assembling, unusual activity recognition, threat treatment, and safe transaction review. Here is a detailed explanation of the processing steps with other formative equations required to achieve total comprehension of the underlying processes involved. Data acquisition and preparation is the first step where data collected from various sources like IoT devices, network logs, and blockchain transactions are procured. Cleaning and preparation of the data are one of the roles of preprocessing to clean up the noise. Normalization is targeting the equality in scaling data and it can be mathematically shown as in (1)

$$x'=x-min(x)/\ max(x)-min(x). \tag{1}$$

Here x ′ is the normalized value and min(x) and max(x) refers to the minimum and maximum value of the dataset respectively. **Feature extraction** followed by, converting raw data into meaningful attributes. For time-series data, statistical features like skewness factor and kurtosis relation are computed for understanding the distribution by the equation (2) & (3)

$$Skewness=((1/N)\ \sum i(xi-\mu)3)/\ \sigma3\ ,(\ i=\ 1\ to\ N). \tag{2}$$

$$Kurtosis=((1/N)(\sum i(xi-\mu)4/\ \sigma4)-3,(\ i=\ 1\ to\ N). \tag{3}$$

These features make it possible to detect outliers that are defined as values which go against the distribution standard. Here xi stands for data, μ for mean and σ for standard deviation. In an anomaly detection phase, data is analyzed by models such as Support Vector Machines SVM to distinguish between normal and anomalous. SVM constructs a hyperplane that is defined by equation (4)

$$f(x)=w\cdot x+b=0. \tag{4}$$

weights: w; features: x; and Bias: b. The gap between classes is then widened as far as possible under the constraints imposed on the cost function in (5)

$$min\ w(1/2)\ \|w\|2+C\sum i\xi i.,(\ i=\ 1\ to\ N). \tag{5}$$

with the restriction made to the constraint in equation (6)

$$yi(w\cdot xi+b)\geq 1-\xi i,\xi i\geq 0. \tag{6}$$

Here, yi is the class label, ξi's are slack variables and C is the penalty parameter. SVM common for complex dataset utilize the kernel functions which include the radial base function (RBF); which maps data into a higher dimensionality space (7)

$$K(xi,xj)=exp(-\gamma\|xi-xj\|2). \tag{7}$$

where γ controls the impact of an individual example. The second element is a real-time monitoring that constantly watches network traffic and the flow of transactions on the block chain identifying any suspicious activity. End/User Communications detected anomaly initiates Response and Mitigation Module, The RL determine the right actions. Similar to (8) the RL framework uses a value function.

$$V(s)=E[\sum tytRt|s0=s]\ ,(\ t=\ 0\ to\ \infty). \tag{8}$$

where V(s) refers to cumulative reward that can be expected from state s, Rt which is the reward at discrete time t and γ represents the discount value. Indeed, an adversarial node or IP address is isolated by modifying the set of coalition's neighbors in the network adjacency matrix A given by (9)

$$A'=A-\Delta A. \tag{9}$$

where *ΔA* erases links to the fallen node. The system also guarantees completeness of blockchain through hashing of transactions for permanent recording. The hashing function H is given in (10)

$$H=SHA\text{-}256(T\|P). \tag{10}$$

Here T is the current transaction, P is the hash of the previous block the ∥ symbol represents concatenation. The Proof of work consensus mechanism is indicated mathematically as (11)

$$H(B)<T. \tag{11}$$

where H(B) represent block hash and T. In order to make the thresholding mechanism capable of providing high detection sensitivity, statistical parameters are employed. Anomalies are detected whenever Z-score is above a certain limit in (12)

$$Z=X-\mu/\ \sigma,flag\ if\ |Z\square>\tau. \tag{12}$$

X: observed value , τ : threshold. The feedback loop of this system optimizes the detection algorithms through learning and reduces the loss function of the specified model. In the case of neural networks the weights are updated using gradient descent employed in (13) backpropagation algorithm.

$$\Delta w=-\eta(\partial J(w)/\partial w). \tag{13}$$

where *Δw* to be the weightages adjustment, *η* as the learning factor, and *J(w)* as the loss factor. While successfully offering real-time cybersecurity solutions, the proposed system is built upon the integrity of blockchain and efficient mathematical modelling incorporated within sound computational architectures. This approach guarantees threats identification and elimination, scale changes in threats, secure transaction log keeping.

Results and discussion

The proposed system promises to add layers of security to the safety envelope by integrating blockchain with machine learning algorithms to detect APTs in distributed networks. Real-time monitoring enhances threat identification compared to traditional batch monitoring, and the anomaly detection algorithm increases its accuracy as well, while blockchain auditing provides traceability. The response of the system is based on reinforcement learning to adapt such new changing attack patterns for efficiency. It has low false positive rates and low computation costs, allowing its scalability for large networks as well as flexibility when faced with different cybersecurity ventures.In summary, system is shown to be a trustworthy and creative tool for combating today's security threats.

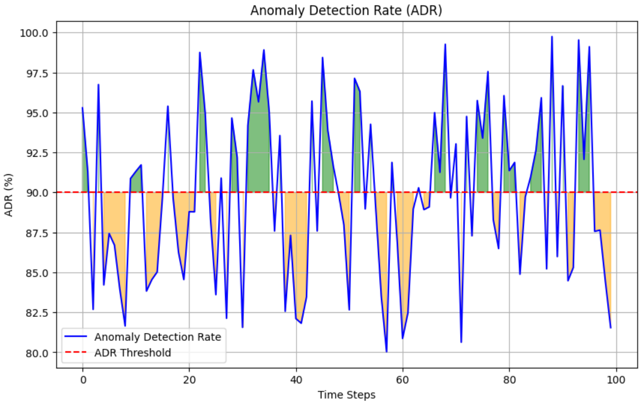

Fig. 3: Threat determining analysis of Proposed framework.

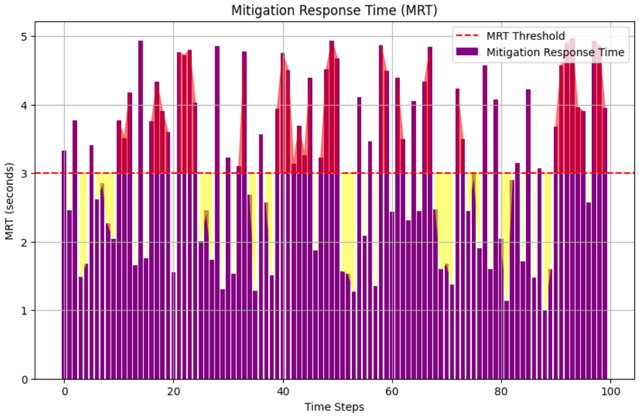

Fig. 4: Mitigation comparison over time spans.

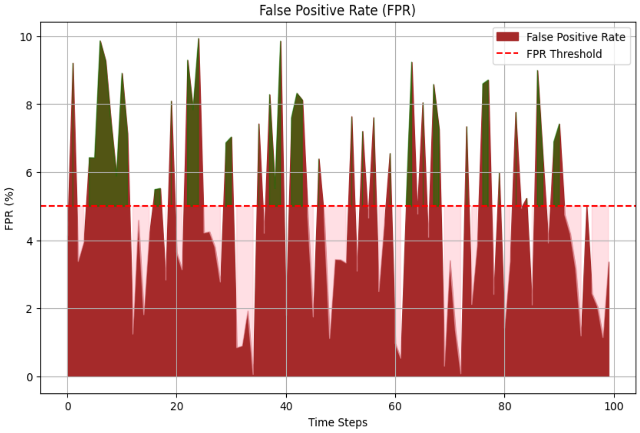

Fig. 5: FPR rate of Proposed framework.

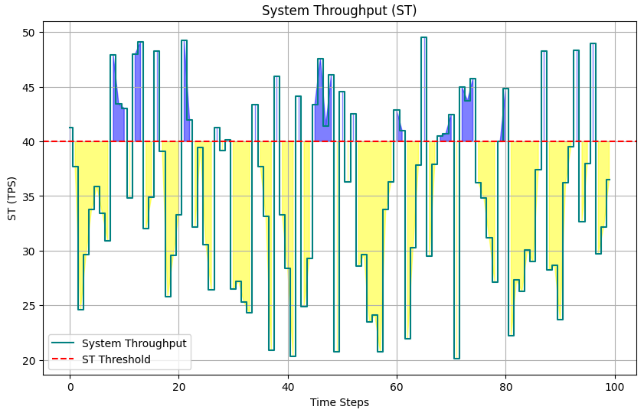

Fig. 6: Overall throughput of Proposed framework.

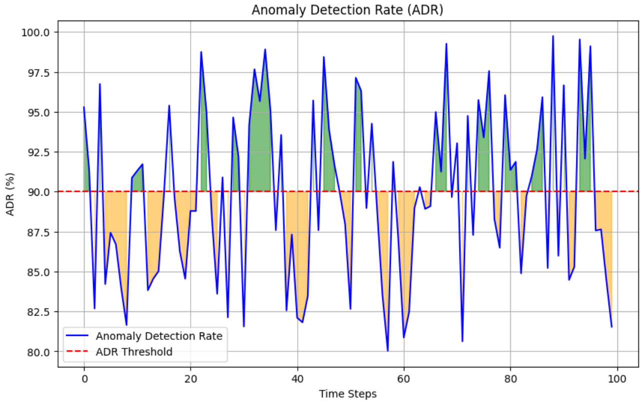

Figure 3 reflects Anomaly Detection Rate (ADR) trends. The level at which ADR is expected to be achieved has been set at 90%.As we note, these values are marked in green once the condition of >90% ADR is met to indicate high detection of relevant information. But if ADR drops below this level, it is represented in orange color pointed at the reduction in system effectiveness. For instance, at time steps 10, and 40, the ADR is greater than 0.9 indicating that there is maximum detection while at time steps 60 & 80, ADR is less than 0.9 indicating that there could be room for system improvement.

In bar plot of figure 4, Mitigation Response Time (MRT) is shown. plot. The mean response time for MRT is set at accepting responses less than 3 seconds with responses colored red implying delay in mitigation. Those that are below or on the threshold are highlighted in yellow representing timely responses. At time steps 5, 30, the MRT is above the threshold level which point out a slow rate of mitigation process, On the other hand at time step 70, the values of MRT is much less than 3 second which is desirable in cases of system performance. The consistent crossing above the threshold rate, like as at steps 20 and 50, showcases areas where optimization process is needed. Figure 5 focuses on the False Positives Rate (FPR) with an area type plot. The threshold value for FPR will set to 5%, and when the values get exceed the threshold rate, it may denoted as marked in green, showcasing a high occurrence of false positives values. Below the threshold value, the value shown in pink marking, representing acceptable rates. For instance, at time slot 15 ms, the FPR get well below the value 5%, suggesting good system accuracy rate, while at time stamp 45, get exceeds 5%, shown in green marking, which denoted a higher count of false positives cases that will leads to unnecessary threats. Such occurrences of higher value of FPR, like at time stamps between 60 and 85, needs attention to improvise system precision factor.

Figure 6 represented the proposed System Throughput (ST) by a step plot. The threshold for ST will set on 40 transactions per second (TPS). When overall throughput exceeds the threshold, the graph showcases values in blue, denoting optimal performance level. When overall throughput falls below this threshold-level, values are denoted in yellow, signaling m8nimized efficacy. For an instance, at time steps 20 and 50, the throughput crossing 40 TPS, indicating effective system operation, while at time stamps between 35 and 75, it falls below range of 40 TPS, suggesting potential performance bottlenecks issues. The significant drops in throughput at time step 90 denotes an area for performance enhancement factor.

Conclusion

In conclusion, the performance analysis of the four key system metrics gives valuable insights into the system's entire performance and showcases areas for potential improvement needed. The Anomaly Detections Rate (ADR) generally remains above rate of 90% threshold values, indicating effective anomaly identification process; however, occasional drops at time value 60 and 80 suggest opportunities for improving identification efficacy. The Mitigation Responses Time (MRT) varies, with time values 5, 30, 20, and 50 crossing the 3-second threshold-level, showcasing to slower response times that will influence entire system performance. On the other hand, when MRT falls below 3 seconds, like at time step 70, the system denoted satisfactory performance rate. The False Positives Rate (FPR) gives instances at time values of 45, 60, and 85 where it exceeded the 5% threshold-level, highlighted increased occurrences of false positives that influenced unnecessary alarms and undermine system accuracy level. Lastly, proposed System Throughput fluctuates over the 40 TPS threshold-level with time varies 20, 50, and 35 representing optimal throughput level, while time stamps 35 and 75 show minimized system throughput, potentially affecting system efficacy. Hence, while the system will performs well, there are specific time steps where enchantment are necessary to ensure consistent and optimal performance over all metrics.

References

1. Pandey, A. K., Saxena, R., Awasthi, A., & Sunil, M. P. (2023). Privacy preserved data sharing using blockchain and support vector machine for industrial IOT applications. Measurement: Sensors, 29, 100891.
2. KHAN, M. Z., BANERJEE, A., & DALAPATI, G. K. Cipher Block Chaining Support Vector Machine for Secured Decentralized Cloud Enabled Intelligent IoT Architecture.
3. Chen, Y. C., Hsu, S. Y., Xie, X., Kumari, S., Kumar, S., Rodrigues, J., & Alzahrani, B. A. (2024). Privacy preserving support vector machine based on federated learning for distributed IoT-enabled data analysis. Computational Intelligence, 40(2), e12636.
4. Singh, K. D. (2021, December). Particle swarm optimization assisted support vector machine based diagnostic system for dengue prediction at the early stage. In 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N) (pp. 844-848). IEEE.
5. Jin, T., Liang, F., Dong, X., & Cao, X. (2023). Research on land resource management integrated with support vector machine—Based on the perspective of green innovation. Resources Policy, 86, 104180.
6. Geetha, C., Johnson, S. D., Oliver, A. S., & Lekha, D. (2024). Adaptive weighted kernel support vector machine-based circle search approach for intrusion detection in IoT environments. Signal, Image and Video Processing, 18(5), 4479-4490.

7. Jagadeesan, J., & Kirupanithi, D. N. (2023). An optimized ensemble support vector machine-based extreme learning model for real-time big data analytics and disaster prediction. Cognitive Computation, 15(6), 2152-2174.
8. Al-Naeem, M., Hafizur Rahman, M. M., Banerjee, A., & Sufian, A. (2023). Support vector machine-based energy efficient management of UAV locations for aerial monitoring of crops over large agriculture lands. Sustainability, 15(8), 6421.
9. Li, J., Li, Y., Song, J., Zhang, J., & Zhang, S. (2024). Quantum support vector machine for classifying noisy data. IEEE Transactions on Computers.
10. Rajasekaran, G., Velavan, P., & Vaidianathan, B. (2024). Least-Squares Support Vector Machine-Based Cancer Prediction System. International Journal of Integrative and Modern Medicine, 2(5), 307-317.
11. Karami, A., & Niaki, S. T. A. (2024). An Online Support Vector Machine Algorithm for Dynamic Social Network Monitoring. Neural Networks, 171, 497-511.
12. Ma, C., Qin, S., Hu, J., & Yan, L. (2021, June). Subway Flow Prediction Based on Improved Support Vector Machine. In 2021 8th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2021 7th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom) (pp. 115-120). IEEE.
13. Arshid, K., Jianbiao, Z., Hussain, I., Lema, G. G., Yaqub, M., & Munir, R. (2024). Support vector machine approach of malicious user identification in cognitive radio networks. Wireless Networks, 30(6), 4761-4772.
14. Navia-Vázquez, A., Díaz-Morales, R., & Fernández-Díaz, M. (2022). Budget distributed support vector machine for non-id federated learning scenarios. ACM Transactions on Intelligent Systems and Technology (TIST), 13(6), 1-25.
15. Hurst, W., Tekinerdogan, B., Alskaif, T., Boddy, A., & Shone, N. (2022). Securing electronic health records against insider-threats: A supervised machine learning approach. Smart Health, 26, 100354.
16. Chander, B. (2022). Artificial Neural Networks and Support Vector Machine for IoT. Artificial Intelligence-based Internet of Things Systems, 77-103.
17. Anitha, P., & Srimathi, C. (2021). Blockchain based Lebesgue interpolated Gaussian secured information sharing for pharma supply chain. International Journal of Intelligent Networks, 2, 204-213.
18. Xiaoqun, L., & Run, L. (2022, May). An improved K-means clustering model based on support vector machine for health insurance cost prediction. In 2022 IEEE 2nd International Conference on Electronic Technology, Communication and Information (ICETCI) (pp. 521-525). IEEE.
19. Tang, W. (2024). Application of support vector machine system introducing multiple submodels in data mining. Systems and Soft Computing, 6, 200096.
20. Bahnam, B. S., & Abd Dawwod, S. (2022). A proposed model for diabetes mellitus classification using coyote optimization algorithm and least squares support vector machine. Int J Artif Intell ISSN, 2252(8938), 1165.