

CyberSentinel-DB: A Machine Learning-Based Framework for Detecting and Preventing Emerging Cyber Threats in Digital Banking

R. Sandhiya¹; Dr. K. Priya²

¹Research Scholar, PG & Research Department of Computer Science, Marudhar Kesari Jain College for Women (Autonomous), Vaniyambadi.

²Research Supervisor, PG & Research Department of Computer Science, Marudhar Kesari Jain College for Women (Autonomous), Vaniyambadi.

Publication Date: 2026/10/21

Abstract: The rapid growth of digital banking services has increased the convenience of online financial transactions. However, it has also exposed banking systems to emerging cyber threats such as phishing, ransomware, Distributed Denial-of-Service (DDoS) attacks, account takeover, and suspicious transaction activities. Traditional rule-based security systems may struggle to identify previously unseen attacks and adapt to changing threat patterns. This study proposes CyberSentinel-DB, a machine learning-based framework for detecting and preventing emerging cyber threats in digital banking. The proposed framework combines supervised machine learning techniques for classifying known cyber threats with unsupervised learning techniques for identifying unusual activities and potential unknown attacks. Transactional patterns, login behavior, IP addresses, transaction frequency, and session-related features can be analyzed to identify suspicious activities. The framework also considers real-time threat monitoring, automated alerts, and preventive actions to improve banking security. Model performance can be evaluated using accuracy, precision, recall, F1-score, false-positive rate, and detection latency. The proposed framework aims to support adaptive, scalable, and privacy-aware cyber threat detection in digital banking environments.

Keywords: *CyberSentinel-DB, Digital Banking Security, Machine Learning, Cyber Threat Detection, Anomaly Detection, Phishing Detection, Ransomware, DDoS Attacks.*

How to Cite: R. Sandhiya; Dr. K. Priya (2026) CyberSentinel-DB: A Machine Learning-Based Framework for Detecting and Preventing Emerging Cyber Threats in Digital Banking. *International Journal of Innovative Science and Research Technology*, 11(10), 1018-1022. <https://doi.org/10.38124/ijisrt/26oct570>

I. INTRODUCTION

The development of digital banking has transformed the way customers access financial services, transfer money, make payments, and manage their accounts. Internet banking, mobile banking applications, and digital payment platforms enable convenient and rapid financial transactions. However, the increasing use of these services has created new opportunities for cybercriminals to exploit vulnerabilities in banking systems.

Cyber threats such as phishing, ransomware, DDoS attacks, credential theft, account takeover, and fraudulent transaction activities can cause financial losses and compromise customer information. Traditional security mechanisms primarily depend on predefined rules and known attack signatures. These mechanisms may not be sufficient to detect new or evolving attacks that do not match existing patterns.

Machine learning provides an opportunity to improve cyber threat detection by identifying patterns in transaction data, network traffic, and user behavior. Supervised learning algorithms can classify known threats, while unsupervised learning techniques can identify anomalies that may indicate emerging attacks.

The proposed CyberSentinel-DB framework integrates these machine learning approaches to strengthen digital banking security. It aims to identify suspicious activities, generate real-time alerts, and support preventive measures while considering data privacy, system performance, and the need for interpretable security decisions.

➤ Objectives

The objectives of the proposed study are:

- To develop a machine learning-based framework for detecting cyber threats in digital banking.

- To identify known cyber attacks using supervised machine learning algorithms.
- To detect unusual transaction patterns and emerging threats using unsupervised learning techniques.
- To analyze transaction frequency, login behavior, IP addresses, and session-related features for identifying suspicious activities.
- To support real-time threat monitoring, alert generation, and preventive security actions.
- To evaluate the framework using accuracy, precision, recall, F1-score, false-positive rate, and detection latency.
- To incorporate privacy-aware data processing and explainable machine learning techniques to improve transparency and trust.

II. LITERATURE REVIEW

Cybersecurity in digital banking has become an important research area due to the increasing number and complexity of cyber attacks. Researchers have explored machine learning techniques to detect malicious network traffic, identify fraudulent transactions, and recognize suspicious user behavior.

Supervised machine learning algorithms, including Logistic Regression, Decision Trees, Random Forest, and Support Vector Machines, are commonly used to classify activities based on labeled data. These algorithms can help distinguish legitimate activities from known malicious patterns.

Unsupervised learning techniques, such as K-Means clustering and Isolation Forest, can identify unusual patterns without requiring every attack type to be labeled in advance. Such methods are useful for investigating suspicious transactions, abnormal login behavior, and deviations from established network activity.

Deep learning techniques, including Convolutional Neural Networks and Recurrent Neural Networks, can also be investigated for analyzing complex patterns in network traffic and sequential transaction behavior. In addition, explainable artificial intelligence methods such as SHAP and LIME can help security analysts understand why a model identifies an activity as suspicious.

Despite these developments, challenges remain in detecting previously unseen attacks, reducing false-positive alerts, processing large volumes of data in real time, and protecting sensitive banking information. CyberSentinel-DB addresses these research challenges through a proposed hybrid framework combining supervised classification, unsupervised anomaly detection, threat monitoring, and privacy-aware processing.

- Note: Specific studies, author names, publication details, and citations should be verified and added before submitting the final research paper.

III. METHODOLOGY

The proposed CyberSentinel-DB framework uses machine learning and behavioral analysis to identify potential cyber threats in digital banking. The methodology consists of dataset preparation, data preprocessing, feature extraction, model development, threat classification, anomaly detection, preventive actions, and performance evaluation.

➤ Dataset Preparation

The study can use publicly available cybersecurity datasets and, where authorized, anonymized banking transaction or login records.

Potential datasets include:

- CICIDS2017 or CICIDS2018 for network intrusion and malicious traffic analysis.
- NSL-KDD for evaluating intrusion detection approaches.
- Suitable financial fraud datasets for suspicious transaction analysis.
- Anonymized digital banking logs, if authorized access is available.

The selected data should contain relevant features such as network traffic statistics, source and destination IP information, connection duration, login attempts, transaction frequency, and activity labels wherever available.

The datasets must be examined for missing values, duplicate records, class imbalance, and suitability for the intended detection task. Network attack detection and financial fraud detection should be evaluated separately where their labels and feature definitions differ.

➤ Data Preprocessing

Data preprocessing prepares raw information for machine learning. The following steps can be applied:

- Remove duplicate records and handle missing values.
- Convert categorical features into numerical representations where required.
- Normalize or standardize numerical features when appropriate.
- Encode known threat labels for supervised classification.
- Handle class imbalance using suitable techniques.
- Separate training and testing data to prevent information leakage.

Sensitive information should be removed, anonymized, or appropriately protected before analysis.

➤ Feature Extraction

Relevant features are extracted from transaction records, authentication logs, and network traffic.

Table 1 Feature Extraction

Feature category	Example features	Purpose
Transaction behavior	Transaction frequency, amount, time interval	Identify unusual transaction patterns
Login behavior	Failed login attempts, login time, session duration	Detect suspicious authentication activity
Network activity	IP address attributes, packet rate, connection duration	Identify potentially malicious network traffic
User behavior	Changes in access patterns, unusual session activity	Identify deviations from normal behavior
Threat indicators	Known attack labels, suspicious connection patterns	Classify recognized cyber threats

The selected features depend on dataset availability and the specific threat being investigated.

➤ *Supervised Machine Learning Classification*

Supervised learning is used to classify known cyber threats using labeled training data.

Potential algorithms include:

- Random Forest: Classifies known threat patterns using multiple decision trees.
- Logistic Regression: Provides a baseline classification model.
- Support Vector Machine (SVM): Separates classes using learned decision boundaries.
- Convolutional Neural Network (CNN): Can be evaluated for suitable structured or transformed data representations.

The model predicts whether an observed activity belongs to a known threat category or a legitimate activity class. The final algorithm should be selected based on experimental comparison rather than assumed performance.

➤ *Unsupervised Anomaly Detection*

Unsupervised learning is used to identify activities that differ significantly from established behavior, including potentially unfamiliar attack patterns.

Potential techniques include:

- Isolation Forest: Identifies observations that are unusual relative to the dataset.
- K-Means Clustering: Groups similar activities and helps identify records that differ from established clusters.
- Autoencoder: Learns normal data patterns and can flag records with unusually high reconstruction errors.

Activities identified as anomalous should be treated as potential threats requiring risk assessment, not automatically assumed to be confirmed attacks. The framework can combine anomaly scores with supervised model predictions to prioritize suspicious events.

➤ *Real-Time Threat Monitoring*

A proposed real-time processing pipeline can use Apache Kafka for event streaming and Apache Spark for data processing.

The workflow consists of:

- Collecting authorized network, authentication, and transaction events.

- Processing and extracting relevant features.
- Applying trained classification and anomaly detection models.
- Generating alerts for suspicious activities.
- Assigning risk levels based on model outputs and security rules.
- Recording the detection results for authorized security review.

The actual processing latency and scalability must be measured during implementation.

➤ *Threat Prevention and Response*

After a suspicious activity is detected, CyberSentinel-DB can support preventive actions according to predefined security policies.

Examples include:

- Generating alerts for security administrators.
- Requesting additional authentication for suspicious login attempts.
- Temporarily restricting high-risk transactions for verification.
- Applying network-level blocking or rate limiting when justified.
- Recording incidents for further investigation and audit.

High-impact actions should use appropriate confidence thresholds, safeguards, and human oversight to reduce unnecessary account restrictions and disruptions to legitimate banking activities.

➤ *Privacy and Explainability*

The proposed framework considers privacy protection and interpretable model outputs as important design requirements.

- Differential privacy: Can be investigated to reduce the disclosure of sensitive information during suitable data analysis or model training.
- Federated learning: Can be explored when multiple institutions need to collaboratively train models without centralizing their raw data.
- SHAP or LIME: Can help explain the features contributing to selected model predictions.
- Access controls and encryption: Can protect sensitive records during storage, processing, and transmission.

The suitability and effectiveness of these techniques must be assessed experimentally and against applicable banking and data-protection requirements.

➤ Performance Evaluation

The proposed framework can be evaluated using the following metrics:

Table 2 Performance Evaluation

Metric	Purpose
Accuracy	Measures the proportion of correctly classified records
Precision	Measures the proportion of predicted threats that are actual threats
Recall	Measures the proportion of actual threats successfully detected
F1-score	Balances precision and recall
False-positive rate	Measures legitimate activities incorrectly flagged as threats
Detection latency	Measures the time required to identify a potential threat

The evaluation should compare individual algorithms with the proposed hybrid framework using consistent data splits and clearly defined threat labels.

- Performance Evaluation and Security Monitoring
- Figure 1. Proposed workflow of the CyberSentinel-DB framework.

➤ Overall Workflow

- Data Collection
- Network traffic, transaction records and login logs
- Data Preprocessing and Feature Extraction
- Supervised Learning
- Known cyber threat classification
- Unsupervised Learning
- Anomaly and unusual activity detection
- Hybrid Threat Assessment
- Combine predictions, anomaly scores and security rules
- Alert Generation and Preventive Response

IV. RESULTS AND DISCUSSION

➤ Performance Evaluation

The CyberSentinel-DB framework can be evaluated using a suitable test dataset containing legitimate activities and different categories of cyber threats. The performance of individual machine learning models and the combined hybrid framework should be compared.

The following table provides a format for recording experimental results after implementation.

Table 3 Performance Evaluation

Performance metric	Result
Accuracy	To be measured
Precision	To be measured
Recall	To be measured
F1-score	To be measured
False-positive rate	To be measured
Average detection latency	To be measured

- Note: These values are intentionally left open because no CyberSentinel-DB experimental results have been provided. Actual results must be obtained through model training and testing.

➤ Classification Analysis

The classification analysis should measure the ability of the supervised learning model to identify known cyber threats correctly. A confusion matrix can be used to examine correct predictions and misclassifications.

The study should specifically examine whether phishing-related indicators, malicious network traffic, suspicious login behavior, and other supported threat categories are detected effectively. Performance should also be reported separately for individual threat classes when sufficient labeled data are available.

➤ Anomaly Detection Analysis

The anomaly detection component should be evaluated using known attack samples, normal activity samples, and suitable previously unseen or held-out threat scenarios.

The analysis should determine whether the framework identifies unusual activities that are not adequately represented in the supervised training data. It should also measure the number of legitimate activities incorrectly flagged as suspicious.

➤ Discussion

The proposed CyberSentinel-DB framework combines supervised classification and unsupervised anomaly detection to address different aspects of digital banking cybersecurity. Supervised learning supports the identification of known threats, while anomaly detection can help identify activities that deviate from established patterns.

Real-time processing may support faster security monitoring, and explainable machine learning can help analysts understand selected model predictions. Privacy-aware processing can further support the responsible handling of sensitive banking information.

However, the framework's practical effectiveness depends on dataset quality, model selection, computational

resources, detection latency, and the ability to distinguish genuine attacks from legitimate but unusual customer behavior. Experimental validation is necessary to establish its accuracy, reliability, and suitability for real-world banking environments.

V. CONCLUSION

This study proposes CyberSentinel-DB, a machine learning-based framework for detecting and preventing emerging cyber threats in digital banking. The framework integrates supervised machine learning for known-threat classification and unsupervised learning for identifying unusual activities. It also incorporates real-time monitoring, automated alert generation, preventive response mechanisms, privacy-aware data processing, and explainable machine learning as key design components.

The proposed approach aims to improve the identification of suspicious activities while reducing the limitations associated with relying exclusively on predefined security rules. Its performance will be evaluated using accuracy, precision, recall, F1-score, false-positive rate, and detection latency.

Future work can focus on implementing the framework using suitable cybersecurity and financial datasets, comparing multiple machine learning algorithms, evaluating previously unseen attack scenarios, and measuring real-time performance. Further investigation into federated learning, differential privacy, and explainable AI may help strengthen privacy, scalability, and transparency in digital banking cybersecurity.

REFERENCES

- [1]. Financial fraud detection using machine learning (2024) L. Hernandez Aros et al., "Financial fraud detection through the application of machine learning techniques: a literature review," *Humanities and Social Sciences Communications*, vol. 11, Article 1130, 2024.
- [2]. Machine learning for financial fraud detection (2022) A. Ali et al., "Financial Fraud Detection Based on Machine Learning: A Systematic Literature Review," *Applied Sciences*, vol. 12, no. 19, 9637, 2022.
- [3]. Financial fraud and anomaly detection (2022) W. Hilal, S. A. Gadsden, and J. Yawney, "Financial Fraud: A Review of Anomaly Detection Techniques and Recent Advances," *Expert Systems with Applications*, vol. 193, Article 116429, 2022.
- [4]. Comprehensive review of financial fraud detection (2021) K. G. Al-Hashedi and P. Magalingam, "Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019," *Computer Science Review*, vol. 40, Article 100402, 2021.
- [5]. Intelligent financial fraud detection (2016) "Intelligent financial fraud detection: A comprehensive review," *Computers & Security*, vol. 57, pp. 47–66, 2016.