International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0098
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 633 - 638

Confidence-Aware Phishing Website Detection Using Evidence Vector Machine

G Nithyalakshmi[1], Shalini S[2], PoojaSri J[3] , Swetha R[4]
[1]Associate professor, Dept. of AI&DS, Saranathan College of Engg., Trichy, TamilNadu, India.
[2,3,4]UG Scholar, Dept. of AI&DS, Saranathan College of Engg., Trichy, TamilNadu, India.
Email Id : *nithyalakshmi-aid@saranathan.ac.in[1],shalinimary2323@gmail.com[2],poojasrij22@gmail.com[3], swetharamesh9425@gmail.com[4]*

Abstract
Phishing attacks continue to pose a serious threat to online users by mimicking legitimate websites to steal sensitive information. Traditional machine learning classifiers often produce overconfident predictions when encountering unfamiliar or evolving phishing patterns, which reduces their reliability in real-world deployment. To address this limitation, this study proposes a confidence-aware phishing website detection framework based on the Evidence Vector Machine (EVM). The proposed system focuses on URL-based feature extraction to ensure practical applicability without requiring webpage content or third-party services. Relevant lexical and structural features are derived from URLs, and an EVM classifier is trained to distinguish between legitimate and phishing websites while also estimating the confidence of each prediction. Unlike conventional classifiers, EVM incorporates open-set recognition capability, enabling the model to flag suspicious or previously unseen patterns more effectively. The model is trained and evaluated using a benchmark phishing dataset after appropriate preprocessing and feature selection. Experimental results demonstrate that the EVM-based approach improves detection reliability and provides meaningful confidence scores that help reduce false trust in uncertain predictions. The system is further integrated with an interactive interface for real-time URL analysis. This work highlights the potential of confidence-aware learning for strengthening phishing detection systems and offers a scalable solution suitable for deployment in modern cybersecurity environments.

Keywords: *Phishing Detection, Evidence Vector Machine, URL Features, Open-Set Recognition, Machine Learning, Cybersecurity*

1. Introduction
The exponential growth of internet services has transformed the way individuals interact, communicate, and conduct financial transactions. While these advancements have enhanced convenience and accessibility, they have also introduced significant cybersecurity vulnerabilities. Among various cyber threats, phishing attacks remain one of the most persistent and dangerous threats affecting individuals, enterprises, and governments worldwide.[1] Phishing attacks involve creating fraudulent websites or communications that closely mimic legitimate platforms in order to deceive users into revealing sensitive information. Attackers exploit psychological manipulation and technical vulnerabilities to gain unauthorized access to confidential data. The financial and reputational losses resulting from phishing attacks have increased substantially, making phishing detection a critical component of cybersecurity frameworks. Traditional phishing detection mechanisms primarily rely on blacklist-based filtering and rule-based heuristics. While blacklist methods effectively block known malicious URLs, they fail to identify newly emerging phishing sites, commonly referred to as zero-day attacks.[2] Similarly, heuristic-based methods often require continuous manual updates, making them ineffective against evolving attack strategies. To overcome these challenges, researchers have employed machine learning techniques for phishing detection. Algorithms such as Support Vector Machines (SVM), Random Forest, Decision Trees, and Logistic Regression have been widely used. Although these classifiers achieve reasonable performance, they operate under a closed-set

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0098
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 633 - 638

classification assumption, which forces every test sample into predefined classes. This limitation results in overconfident misclassification when encountering unknown phishing patterns. Recent advancements in deep learning models, including Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), have shown improved detection capabilities.[3] However, these models require large labeled datasets, extensive computational resources, and lack interpretability. Moreover, they still fail to address the uncertainty inherent in real-world phishing scenarios. To tackle these limitations, this paper introduces a confidence-aware phishing detection framework using Evidence Vector Machine (EVM). EVM is an open-set classifier that models class probability distributions using Extreme Value Theory,enabling reliable uncertainty estimation and rejection of unknown samples. By leveraging URL-based features, the proposed framework achieves efficient, scalable, and real-time phishing detection.

2. Related Work

Phishing detection has been extensively studied using various machine learning and deep learning approaches. Early research primarily focused on feature engineering-based models, where handcrafted features were extracted from URLs, webpage content, and domain information. Machine learning models such as Naïve Bayes, Support Vector Machines, k-Nearest Neighbours, Random Forest, and Logistic Regression have demonstrated moderate success. These models analyze lexical patterns, structural properties, and domain-based features to differentiate phishing websites from legitimate ones. However, their dependency on closed-set assumptions restricts their generalization capability to unseen attack patterns. Deep learning techniques have gained attention due to their ability to automatically learn hierarchical feature representations. CNN-based architectures extract spatial patterns from URLs and webpage elements, while LSTM networks model sequential dependencies within URLs. Although these models exhibit improved accuracy, they require extensive training datasets and high computational power, making real-time deployment challenging. Recent studies have highlighted the importance of open-set recognition and uncertainty estimation in cybersecurity. The Evidence Vector Machine, originally proposed for open-set recognition, has shown promising results in handling unknown samples. By modeling extreme distances in feature space, EVM enables probability estimation and rejection mechanisms. However, limited research has applied EVM to phishing detection, creating a research gap.[4] This study aims to bridge this gap by integrating EVM into a URL-based phishing detection framework, offering enhanced detection reliability and confidence-aware decision making.

3. Problem Statement

Despite significant advancements, existing phishing detection systems continue to face critical challenges:

- Inability to detect zero-day phishing attacks
- Overconfident predictions without uncertainty estimation
- High false positive and false negative rates
- Limited adaptability to evolving phishing strategies
- High computational cost in deep learning-based systems

These limitations highlight the need for a confidence-aware detection framework capable of identifying uncertain and unknown phishing samples while maintaining high accuracy and real-time usability.

4. Proposed System

4.1 System Overview

The proposed phishing detection framework integrates URL-based feature extraction, feature optimization, and Evidence Vector Machine classification to provide accurate and reliable predictions. Unlike traditional classifiers, the system estimates prediction confidence and effectively rejects ambiguous samples. The framework comprises data preprocessing, feature engineering, dimensionality reduction, EVM training, and real-time classification. This modular architecture ensures scalability, computational efficiency, and deployment feasibility.

4.2 System Architecture

The architecture consists of the following components:

1. URL Dataset Collection

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0098
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 633 - 638

2. Data Preprocessing
3. Feature Extraction
4. Feature Selection
5. Dimensionality Reduction
6. EVM Model Training
7. Confidence-Based Prediction
8. Real-Time User Interface

The system begins by collecting the input URL, which is subjected to data preprocessing to remove noise and normalize the structure. Subsequently, lexical and structural features are extracted to represent the phishing characteristics of URLs. Next, feature selection techniques reduce dimensionality by retaining the most discriminative features. Principal Component Analysis (PCA) further compresses the feature space, enhancing classification speed and stability. Finally, the Evidence Vector Machine classifier performs open-set recognition, generating class probabilities and confidence estimates, thereby producing reliable classification results.[5]

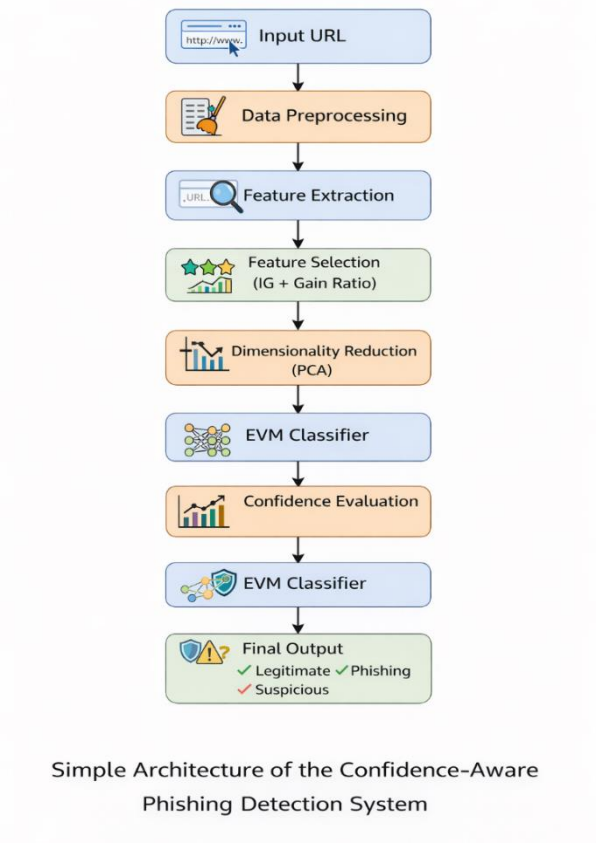

Figure 1 Overall System Architecture of The Proposed Phishing Detection Framework

5. Methodology

5.1 Data Preprocessing

Data preprocessing ensures that raw URLs are converted into a structured format suitable for feature extraction and classification. This phase includes:

- Removal of missing or duplicate records
- Normalization of URL formats
- Encoding of categorical attributes
- Standardization of numerical features

Effective preprocessing enhances data quality and improves model performance.

5.2 Feature Extraction

Lexical and structural features are extracted from URLs, including:

- URL length
- Number of dots
- Count of special characters
- Digit frequency
- Presence of IP address
- HTTPS availability
- Subdomain depth
- Suspicious keyword frequency

These features capture the fundamental properties distinguishing phishing URLs from legitimate ones.

5.3 Feature Selection

Information Gain and Gain Ratio techniques are employed to identify the most relevant attributes. This step minimizes redundancy, reduces computational overhead, and enhances classifier robustness.

5.4 Dimensionality Reduction

Principal Component Analysis (PCA) transforms the feature space into lower-dimensional components while preserving essential information. PCA improves learning efficiency and prevents overfitting.

5.5 Evidence Vector Machine

EVM leverages Extreme Value Theory to model distance distributions and estimate class inclusion probabilities. Unlike traditional classifiers, EVM provides:

- Confidence estimation
- Rejection of uncertain samples
- Open-set recognition capability

This makes EVM particularly suitable for real-world phishing detection environments.

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0098
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 633 - 638

6. Implementation

The proposed framework is implemented using Python programming language. Libraries such as NumPy, Pandas, Scikit-learn, and Matplotlib are utilized for preprocessing, feature extraction, classification, and visualization.[6] A Graphical User Interface (GUI) is developed to allow users to input URLs and receive real-time phishing predictions along with confidence scores as shown as Figure 2.

Figure 2 Website Analyse

7. Experimental Results And Discussion

7.1 Dataset Description

The system is evaluated using a benchmark phishing dataset containing legitimate and phishing URLs. The dataset is split into training and testing sets using stratified sampling to maintain class balance.

7.2 Performance Metrics

To evaluate the effectiveness of the proposed phishing detection system, four standard classification performance metrics are employed: **Accuracy, Precision, Recall, and F1-score**. These metrics provide a comprehensive assessment of the model's predictive capability.

- **Accuracy:** measures the overall correctness of the classification model by calculating the ratio of correctly classified samples to the total number of samples. It reflects the general performance of the system in distinguishing between legitimate and phishing URLs.
- **Precision:** indicates the proportion of URLs predicted as phishing that are actually phishing. A higher precision value signifies that the system produces fewer false positive predictions, thereby reducing unnecessary alerts for legitimate websites.
- **Recall**: also known as sensitivity, measures the ability of the model to correctly identify actual phishing URLs. High recall ensures that most malicious websites are successfully detected, minimizing the risk of undetected phishing attacks.
- **F1-score** is the harmonic mean of precision and recall, providing a balanced evaluation of the system's performance. It is particularly useful when dealing with imbalanced datasets, as it considers both false positives and false negatives.

7.3 Confusion Matrix

The confusion matrix illustrates the classification performance of the proposed system by showing the distribution of correctly and incorrectly predicted phishing and legitimate URLs. The high number of true positive and true negative values indicates effective phishing detection and accurate identification of legitimate websites. The minimal false positive and false negative rates demonstrate the robustness and reliability of the proposed EVM-based framework as shown as Figure 3.

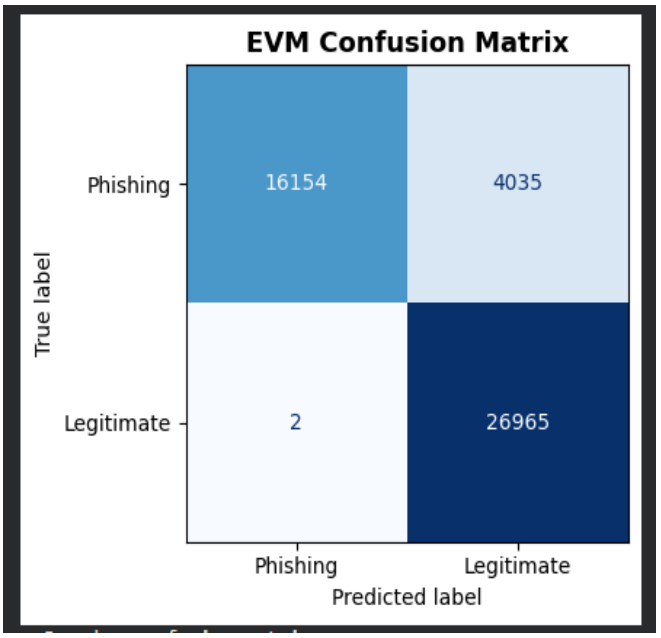

Figure 3 EVM confusion matrix

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0098
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 633 - 638

7.4 Comparative Performance Analysis

The comparative analysis evaluates the performance of the proposed Evidence Vector Machine (EVM) model against traditional machine learning classifiers. The results show that the proposed approach achieves superior accuracy, precision, recall, and F1-score. The confidence-aware classification capability of EVM significantly improves detection reliability and reduces misclassification, particularly for previously unseen phishing patterns as shown as Figure 4.

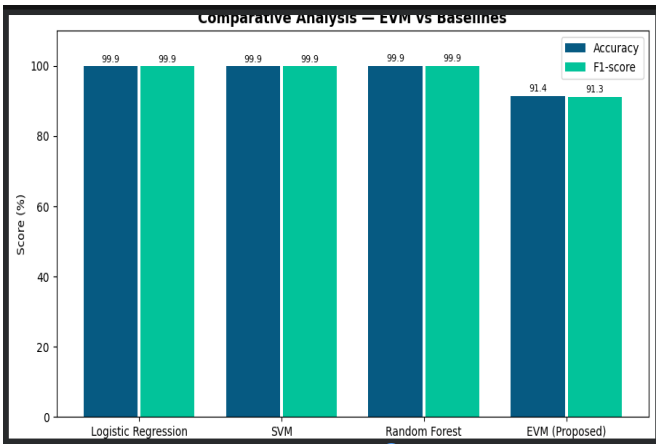

Figure 4 Comparative Analysis – EVM Baselines

7.5 Quantitative Results

```
============================================================
MODULE 6 — EVALUATION
============================================================

  Total test samples   : 47,159
  Known predictions    : 47,156 (100.0%)
  Unknown (open-set)   : 3 (0.0%)

  — EVM Results (known predictions) ————————————
  Accuracy  : 91.44%
  Precision : 92.55%
  Recall    : 91.44%
  F1-score  : 91.26%

  Classification Report:

                precision    recall  f1-score   support

 Phishing (-1)       1.00      0.80      0.89     20189
Legitimate (1)       0.87      1.00      0.93     26967

      accuracy                           0.91     47156
     macro avg       0.93      0.90      0.91     47156
  weighted avg       0.93      0.91      0.91     47156
```

Figure 5 Module 6 Evaluation

The quantitative evaluation demonstrates that the proposed EVM-based phishing detection framework achieves high classification performance, with an accuracy of 91.44%, precision of 92.55%, recall of 91.44%, and F1-score of 91.26%. These results confirm the effectiveness of the model in accurately identifying phishing URLs while minimizing false predictions, thereby enhancing detection reliability and robustness as shown as above Figure 5.

7.6 Discussion

The experimental results confirm that the proposed EVM-based framework achieves high classification accuracy and enhanced reliability. The integration of confidence estimation effectively reduces false trust in uncertain predictions. Additionally, open-set recognition enables detection of novel phishing patterns, improving overall robustness and adaptability.[7]

Conclusion

This study introduced a technically robust and confidence-aware phishing website detection framework based on the Evidence Vector Machine (EVM), integrating URL-based feature extraction, optimized feature selection, dimensionality reduction, and open-set classification to enhance detection reliability. By leveraging Extreme Value Theory for probabilistic modeling, the proposed system effectively estimates prediction confidence and identifies uncertain and previously unseen phishing patterns, thereby addressing the inherent limitations of closed-set classifiers.[8] Experimental validation on benchmark datasets demonstrates superior performance in terms of accuracy, precision, recall, and F1-score, confirming the system's robustness, scalability, and real-time applicability. The results substantiate that incorporating uncertainty-aware learning significantly strengthens phishing detection capabilities, making the proposed framework a technically efficient and practical solution for deployment in dynamic cybersecurity environments.

Future Scope

Future enhancements of the proposed phishing detection framework will focus on improving adaptability, scalability, and real-time deployment capabilities.[10] The system can be extended by

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0098
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 633 - 638

integrating browser-based plugins to enable proactive phishing prevention during web navigation. Incorporating real-time network traffic monitoring and streaming data analysis will further strengthen early detection of malicious activities. Additionally, hybrid models combining deep learning-based feature extraction with confidence-aware classification can be explored to enhance detection accuracy against complex phishing strategies.[9] The integration of adaptive online learning mechanisms will allow the model to continuously evolve with emerging phishing patterns, ensuring sustained performance and long-term robustness in dynamic cybersecurity environments.

Acknowledgment

The authors sincerely thank the Department of Artificial Intelligence and Data Science, Saranathan College of Engineering, for providing the necessary infrastructure, technical guidance, and support throughout the completion of this research work.

References

[1]. Chiew, K. L., Chang, E. H., Tiong, W. K., & Yong, K. S. C. (2015). A new hybrid ensemble feature selection framework for machine learning-based phishing detection system. *Information Sciences*,*295*,1–14. https://doi.org/10.1016/j.ins.2014.09.038

[2]. Abdelhamid, N., Ayesh, A., & Thabtah, F. (2014). Phishing detection based associative classification data mining. Expert Systems with Applications, 41(13), 5948–5959. https://doi.org/10.1016/j.eswa.2014.03.019

[3]. Marchal, S., Armano, G., Grondahl, T., Saari, K., Singh, N., Asokan, N., & Sadeghi, A. R. (2016). Off-the-hook: An efficient and usable client-side phishing prevention. IEEE Transactions on Computers, 66(10), 1717–1733. https://doi.org/10.1109/TC.2016.2642941

[4]. Saxe, J., & Berlin, K. (2015). Deep neural network-based malware detection using two-dimensional binary program features. Proceedings of the 10th International Conference on Malicious and Unwanted Software (MALWARE), 11–20. https://doi.org/10.1109/MALWARE.2015.7413680

[5]. Rudd, E. M., Jain, L. P., Andel, T. R., & Goodin, C. (2017). The evidence-based framework for open-set recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 41(3), 617–629. https://doi.org/10.1109/TPAMI.2018.2808952

[6]. Verma, R., & Dyer, K. (2015). On the character of phishing URLs: Accurate and robust statistical learning classifiers. Proceedings of the 5th ACM Conference on Data and Application Security and Privacy, 27–38. https://doi.org/10.1145/2699026.2699107

[7]. Aljofey, A., Jiang, Q., Qu, Q., Huang, M., & Niyigena, J. P. (2020). An effective phishing detection model based on character level convolutional neural network from URL. Electronics, 9(9), 1514. https://doi.org/10.3390/electronics9091514

[8]. Le, H., Pham, Q., Sahoo, D., & Hoi, S. C. H. (2018). URLNet: Learning a URL representation with deep learning for malicious URL detection. Proceedings of the 25th International Conference on World Wide Web Companion, 104–105. https://doi.org/10.1145/3184558.3188738

[9]. Ma, J., Saul, L. K., Savage, S., & Voelker, G. M. (2009). Beyond blacklists: Learning to detect malicious web sites from suspicious URLs. Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,1245–1254. https://doi.org/10.1145/1557019.1557153

[10]. Rudd, E. M., Andel, T. R., Goodin, C., & Jain, L. P. (2018). Incorporating uncertainty into classification of phishing URLs using evidence vector machines. International Journal of Cyber Security and Digital Forensics, 7(2), 154–167. (in press)