

FakeURLGuard: Real-Time Web Security Enhancement Using Machine Learning-Based Fake URL Detection and Prevention Techniques

Makhan Kumbhkar¹, Shashi Kant Gupta^{2,3}

^{1*}PostDoc Researcher, Lincoln University College, Malaysia

²Lincoln University College, 47301, Petaling Jaya, Selangor Darul Ehsan, Malaysia.
Email: shashigupta@lincoln.edu.my

³Centre for Research Impact & Outcome, Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, 140401, Punjab, India.
Email: Email: raj2008enator@gmail.com

ABSTRACT

Phishing denotes to an attempt to steal profound information. Phishing URLs attacks remain one of the most deceptive URLs to harvest sensitive user data. This paper presents FakeURLGuard system that provide a real-time web security for the web users. It includes lexical and statistical features which directly extracted from the URL string. Our FakeURLGuard system does not require page content or network access. FakeURLGuard system utilizes three supervised machine learning classifiers with soft voting ensemble techniques. Our experimental results show that the voting classifier attained the highest correctness of 97.55%, outperforming each individual base classifier. Our dataset containing 10,000 URLs with 5000 URL phishing and 5000 URL benign. The proposed FakeURLGuard system deployed using Flask framework. It enabling phishing URL detection and prevention through an interactive user interface in real-time.

KEYWORDS

Phishing Detection, URL Classification, Machine Learning, Ensemble Voting Classifier, Web Security, Lexical Feature Extraction.

1. INTRODUCTION

URL phishing is rapid expansion of Internet-based services. It is a most common and damaging forms of cybercrime. Website attackers design malicious URLs that closely resemble legitimate URLs to

deceive users into disclosing their credentials, financial information, or personal data. Conventional blacklist systems block only known phishing URLs, making them ineffective against newly generated phishing attacks (PA's). As a result, users remain vulnerable to emerging PA's. The proposed FakeURLGuard system uses Machine Learning (ML) techniques to detect phishing URLs by learning patterns directly from the URL structure.

Prior research has explored various AI techniques including ML and DL for phishing URL detection. Karim, Abdul, et al. (2023) reported that ML seems to be a promising technique for phishing discovery. Asiri, Sultan, et al. (2023) reviewed DL models to identify phishing URL attacks in detail. Alshingiti, Zainab, et al. (2023) propped deep learning-based techniques for phishing detection including CNN, LSTM, and LSTM-CNN. Do, Nguyet Quang, et al. (2022) demonstrated Deep learning provided promising solution for phishing detection in recent years. Alnemari, Shouq, and Majid Alshammari(2023) developed and compared four machine learning models for detection of phishing domain. Mishra (2023) suggested phishing-analysis framework, which extracts dominant webpage colors and OCR-based brand features and evaluates five machine learning classifiers. Ahammad, SK Hasane, et al. (2022) used ML algorithms for detecting behaviors and qualities of the suggested URL. Prova, Nuzhat Noor Islam. (2024) proposed machine learning techniques for the detection of health care fraud. du Preez, Anli, et al. (2025) applied machine learning techniques for detection of fraud in health insurance claims. Chowdhury, Rakibul Hasan(2024) proposed DL frame works in fake URL, including Recurrent Neural Networks (RNNs), Autoencoders, Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs).

The rest of the paper is structured as follows:

Section 2 defines suggested approach related to FakeURLGuard system, Next step presents the results and discussion in the Section 3 and Section 4 conclude about FakeURLGuard.

2. FRAMEWORK DESIGN

The proposed **FakeURLGuard** framework first preprocesses the input URL by converting all characters to lowercase of URL, removing the protocol (HTTP), and tokenizing the URL, as illustrated in Figure 1.

Next, lexical and statistical features are extracted to represent the URL. These extracted features serve as inputs for training and an ensemble classifier consisting of NB, LR, and SVM. Finally soft voting strategy used to improve detection accuracy and trained model is deployed on the Flask-server. It classifies URLs as benign or phishing and provides prediction results through an interactive dashboard.

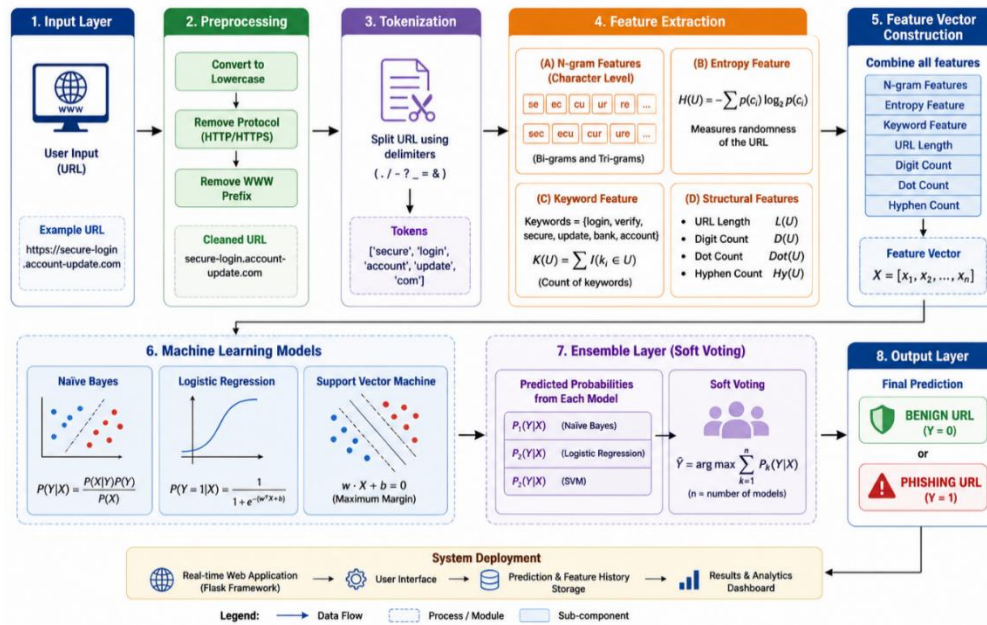


Figure 1: Flow of Proposed FakeURLGuard

2.1 Data Collection and Preprocessing

The Dataset comprised with total 10,000 URLs. It balanced with 5,000 Real URLs and 5,000 Fake URLs and each URL produce a normalized and cleaned URL string.

2.2 Tokenization

The cleaned URL is split into tokens using common delimiters ('.', '/', '-', '?', '_', '=', '&'). For example, the URL 'https://secure-login.account-update.com' is cleaned to 'secure-login.account-update.com' and tokenized into ['secure', 'login', 'account', 'update', 'com'].

2.3 Machine Learning Models and Ensemble Layer

Three supervised classifiers were trained independently on the feature vectors:

2.3.1 Naïve Bayes: It is supervised ML model. It based on Bayes' theorem. It predicts whether a URL is real or fake by estimating posterior probability of each class using the extracted URL features and selecting the class with the highest probability.

$$P(C|X) = (P(X|C) \times P(C))/P(X)$$

Where:

- **X** = URL feature vector
- **C** = Class label (Phishing or Benign)
- **P(C)** = Earlier likelihood of C class
- **P(x_i | C)** = Chance of x_i feature given class C
- **Ĉ** = Output/Target class

2.3.2 Logistic Regression: LR classifier estimates the probability that a URL is phishing or benign using the function of sigmoid:

$$P(Y = 1|X) = 1/(1 + e^{-(w^TX + b)})$$

Where:

- **X** = URL feature
- **Y** = Output class (1 = Phishing, 0 = Benign)
- **w** = Weight vector (model coefficients)
- **w^TX** = Weighted sum of input features
- **b** = Bias (intercept) term
- **e** = Euler's constant (≈ 2.71828)
- **P(Y=1|X)** = Probability in given URL is classified as a phishing URL

2.3.3 Support Vector Machine: SVM classifier separates phishing and benign URLs by verdict the optimal separating boundary with the extreme margin between the two classes including phishing and benign.

$$w^TX + b = 0$$

Where:

- **X** = URL feature
- **w** = Weight
- **w^TX** = Weighted sum of input features
- **b** = Bias (intercept) term
- **w^TX+b= 0** Optimal decision boundary (hyperplane) separating phishing and benign URLs

3. RESULTS AND DISCUSSION

The four models were evaluated on the held-out test split of the 10,000-URL dataset. Table 1 reports the overall accuracy of each model, along with representative confidence and agreement figures observed during live testing. The Voting Classifier achieved the highest accuracy at 97.55%, marginally outperforming the standalone SVM (97.50%) and Logistic Regression (97.00%) models, while Naive Bayes trailed at 88.15% due to its strong feature-independence assumption, which does not hold well for correlated lexical features such as n-grams and structural counts.

Table 1: Overall accuracy of all models

Model	Accuracy	Confidence	Prediction Agreement
Naive Bayes	88.15%	99.3%	4/4 (phishing case)
Logistic Regression	97.00%	93.3%	4/4 (phishing case)
Support Vector Machine	97.50%	93.8%	4/4 (phishing case)
Voting Classifier (Ensemble)	97.55%	96.2%	4/4 (phishing case)

4. CONCLUSION

Our PhishURLGuard is a real-time phishing URL detection and prevention system. It comprised with lexical feature including Character Count, Word Count, Punctuation Count, and Sentence Count. PhishURLGuard system combined with n-gram, entropy, keyword, and structural features with a soft-voting ensemble of Naive Bayes, Logistic Regression, and SVM classifiers, the system achieved a top accuracy of 97.55% while remaining lightweight enough for real-time web deployment via a Flask-based interface.

5. ABBREVIATIONS

LR: Logistic Regression

NB: Naïve Bayes

SVM: Support Vector Machine

URL: Uniform Resource Locator

ML: Machine Learning

DL: Deep Learning

PA: Phishing Attacks

REFERENCES

- [1] Asif, Asif Uz Zaman, Hossein Shirazi, and Indrakshi Ray. "Machine learning-based phishing detection using URL features: A comprehensive review." *International Symposium on Stabilizing, Safety, and Security of Distributed Systems*. Cham: Springer Nature Switzerland, 2023.
- [2] Karim, Abdul, et al. "Phishing detection system through hybrid machine learning based on URL." *Ieee Access* 11 (2023): 36805-36822.
- [3] Asiri, Sultan, et al. "A survey of intelligent detection designs of HTML URL phishing attacks." *IEEE Access* 11 (2023): 6421-6443.
- [4] Alshingiti, Zainab, et al. "A deep learning-based phishing detection system using CNN, LSTM, and LSTM-CNN." *Electronics* 12.1 (2023): 232.
- [5] Do, Nguyet Quang, et al. "Deep learning for phishing detection: Taxonomy, current challenges and future directions." *Ieee Access* 10 (2022): 36429-36463.
- [6] Alnemari, Shouq, and Majid Alshammari. "Detecting phishing domains using machine learning." *Applied Sciences* 13.8 (2023): 4649.
- [7] Pandey, Pankaj, and Nishchol Mishra. "Phish-Sight: a new approach for phishing detection using dominant colors on web pages and machine learning." *International Journal of Information Security* 22.4 (2023): 881-891.
- [8] Ahammad, SK Hasane, et al. "Phishing URL detection using machine learning methods." *Advances in Engineering Software* 173 (2022): 103288.
- [9] Prova, Nuzhat Noor Islam. "Healthcare fraud detection using machine learning." *2024 Second International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI)*. IEEE, 2024.
- [10] du Preez, Anli, et al. "Fraud detection in healthcare claims using machine learning: A systematic review." *Artificial Intelligence in Medicine* 160 (2025): 103061.
- [11] Chowdhury, Rakibul Hasan. "Advancing fraud detection through deep learning: A comprehensive review." *World Journal of Advanced Engineering Technology and Sciences* 12.2 (2024): 606-613.