Computing and Informatics, Vol. 45, 2026, 382–409, doi: 10.31577/cai_2026_2_382

PREDICTING EMERGING PHISHING THREATS THROUGH LIGHTWEIGHT AI DATA MINING ON THE EDGE

Ahmad Mohammed Almorabea, Usman Ali Khan, Muhammad Binsawad, Syed Hamid Hasan*

Department of Information Systems
Faculty of Computing and Information Technology
King Abdulaziz University
Jeddah, Saudi Arabia
e-mail: aelmerabea@stu.kau.edu.sa, {ukhan, mbinsawad, shhasan}@kau.edu.sa

Abstract. Phishing remains a cybersecurity risk by exploiting human nature and technology loopholes to steal sensitive data. In this research, with secondary qualitative analysis including academic databases and content analysis, machine-learning-based phishing models with multidimensional feature analysis are critically analyzed. Computational inefficiency, real-time detection capability, and ethics concerns like model explainability and privacy are some of the challenges. Model reliability is the primary issue of feature ablation analysis of features like URL structure, domain registration, and webpage behavior as well as comparative performance evaluation of classifiers like Naive Bayes, SVM, Random Forest, Decision Trees, and Gradient Boosting. Scalable, adaptive, and ethical phishing is proposed using NLP and blockchain-based hybrid solutions for long-term digital pathology systems.

Keywords: Phishing detection, machine learning, predictive models, cybersecurity, feature analysis, hybrid models, natural language processing, blockchain, adaptive frameworks

Mathematics Subject Classification 2010: 68T05

* Corresponding author

1 INTRODUCTION

Phishing is the most prevalent and covert threat in the cyber world [1]. Phishing uses social engineering to deceive people into divulging private information like banking details, login credentials, and identities. Despite the technological push to maximize cybersecurity, phishing attacks continue and entry points of system compromise are increasingly reliant on human weaknesses. Today, these attacks are more sophisticated because they use visually deceiving websites, seemingly legitimate domains, and secure connection using SSL/TLS certificates. This dynamic only goes to highlight the need for more intelligent and proactive detection systems that possess the ability to identify current phishing attempts and stop them before they can even occur.

Phishing is much more dangerous in the healthcare environment, where patient safety can be directly threatened by hacking of EHRs, lab reports, and digital pathology images. Telepathology's increasing popularity and AI-driven diagnostic platforms becoming the new norm, makes medical personnel the focus of attackers with spoofed hospital portal websites, laboratory access URLs, and credential-harvesting email campaigns. Predictive phishing detection in health is therefore not just a cybersecurity issue, but also imperative to building reliable, resilient, and sustainable health systems.

The fast digital revolution in healthcare such as telepathology expansion and confidential diagnostic images exchange raised the exposure to phishing attacks, as proposed by emerging studies highlighted within this paper. Past research tends to commonly discuss frequent phishing detection methods without recognizing the specific challenges in digital pathology and healthcare information infrastructures [2]. Rule-based policy regulations for requiring confidentiality for patient data like HIPAA and GDPR call for detection methods that aim to maintain a balance between correctness, interpretability, and compliance. These are conjoint factors that form a pressing need for efficient and ethically sound predictive models that are intended to meet the particular purposes of sustainable healthcare systems.

The main objectives of this research are as follows:

- To comprehensively investigate and compare a wide range of machine learning classifiers for phishing detection in sustainable healthcare systems, with an emphasis on digital pathology workflows.
- To extend beyond standard models (Naive Bayes and SVM) to cover and compare other algorithms such as Random Forest, Decision Trees, and Gradient Boosting Machines.
- To carry out an ablation study that rationally isolates and compares the importance of different groups of features in specific, URL structure, domain registration data, and page actions on phishing detection performance.
- To address the unique challenges of healthcare phishing detection by developing predictive models that are not computationally intensive and can be deployed efficiently on edge devices while maintaining high detection accuracy.

While NB and SVM are proven baseline approaches to phishing detection, their application as the sole choice inhibits comparative assessment resilience. To counter this, our work incorporates Random Forest (RF), Decision Trees (DT), and Gradient Boosting Machines (GBM) as well as NB and SVM. An ablation study is also presented that disables or separates feature sets (URL anatomy, domain information, webpage behavior) systematically to highlight their contribution in the performance of phishing detection. This double approach supports the conclusions by giving both model-level and feature-level outcomes.

2 LITERATURE REVIEW

Previous literature on phishing detection tended to focus on heuristic and list-based mechanisms, which were sufficient when the main threat came from static attacks. Consequently, contrary to expectations, the use of traditional approaches diminishes, as Olivo et al. claimed [3]. In the same regard, Chiew et al. [4] pointed out that heuristic approaches are bound to be restricted by reliance on some patterns, given that the modus operandi of the distinct class of attackers consists of incorporating minor alterations to their methodology. These gaps have boosted the development of modern machine-learning-based models that possess the capacity to conform to, update, and incorporate exploitable data based on new phishing techniques regarding previous data that can be analysed to flag signs of wrongdoing [5].

Machine learning is proven to have a high potential for phishing detection, with especially good data processing and recognising subtle features of communication submissions regarding phishing or not. In their study, Orunsolu et al. [6] highlight that support vector machines (SVMs) and naive Bayes machine learning models can attain extremely high accuracy rates if efficient feature selection methods are applied in order to build efficient predictive models. Instead, they use some data, such as URL structure, domain registration details, and web page behaviour, and by doing so they can detect phishing with high precision. However, as pointed out by Tang and Mahmoud [7], the efficacy of these models is only guaranteed by using a feature vector of high quality and completeness, and therefore, the extraction of rich features is of immense importance.

In addition to examining traditional classifiers such as Naive Bayes and SVM, existing literature has widened comparative studies to include more complex machine learning and deep learning techniques [5]. Some of the main objectives of this review include identifying the evolution of phishing detection approaches, comparing their relative efficacy, and summarizing existing trends and areas of future work. In addition to basic algorithms, studies now entail ensemble methods, hybrid models, and more sophisticated approaches like CNN, RNN, and other deep architectures [8]. Comparative studies have evaluated these approaches in terms of detection efficacy, adaptability in the face of new and innovative phishing techniques, computational expense, and practicability for real-world applications. This wider focus allows a deeper exploration of the pros and cons of each approach and

uncovers valuable avenues for future research and development in phishing detection.

Many machine-learning-based systems take the detection of phishing as a binary classification problem, where given a site, it can classify it as phishing or legitimate. Chy [9] explains that features of supervised learning techniques, such as logistic regression and decision trees, enable them to perform well when classifying data based on specific characteristics. For instance, it has been proven that logistic regression is good at identifying statistical relations but bad at understanding features like URL length and domain age, while decision trees are good at visualising the classification criteria. Nevertheless, the possible overfitting and computational overhead from large feature sets pose challenges that motivate the design of optimised models that strive to meet accuracy with efficiency [10, 11].

Machine-learning-based phishing detection systems have their own disadvantages. Supervised learning methods, for instance, are based on labelled datasets, meaning that they are also unlikely to deal with the new or previously unseen phishing techniques. As was indicated by Moghimi and Varjani [12], phishing datasets are unbalanced and can adversely impact the model by a high false positive or negative rate. There are methods towards learning models in a step of unsupervised learning through which the model may be able to recognise outliers away from the normal patterns, and this may be helpful in this regard. However, such models may not be as accurate as the supervised models or may require some fine-tuning to be able to work properly [13].

It is also worth mentioning that such a procedure has been designed to include the use of new and state-of-the-art technologies, such as deep learning and NLP, as an advanced form of phishing detection. Due to its ability to work with high-dimensional data, deep learning models have been observed to have the power of learning features that relate to phishing. It was demonstrated additionally by Yang et al. [14] that the use of multidimensional characteristics obtained by means of deep learning significantly improves detection functions, in particular, in real-time systems. Similarly, NLP methods allow for processing the textual message from the received emails or from the web page, finding out the peculiarities of the used language and grammar and semantic peculiarities indicating the presence of phishing emails. Nevertheless, the general reluctance to brave such an undertaking is justified, due to the computational demands and limited interpretability associated with these advanced models, as highlighted by Tang and Mahmoud [7].

The first problem that phishing detection has to deal with is the almost infinite number of techniques that attackers have and can use to launch their attacks. For instance, domain spoofing consists of generating URLs that look similar to trustworthy ones, for example, modifying the URL by character or creating subdomains [9]. According to Alabdan [15], SSL/TLS certificates, which used to be a sign of a secure website has been used to add legitimacy to phishers' fraudulent sites. Visual similarity techniques, those that replicate the design and layout of trusted websites make the detection more difficult. Finally, in phishing detection, the interactions of these techniques highlight that phishing implies a multi-sided effort by considering

features from many angles to depict a full range of varieties in phishing methodology [5].

The impact of phishing extends beyond the economic loss but involves reputational damages, breaks in business activities, and legal issues for the organisations that have been involved. According to Rao et al. [16], the loss of confidential data not only results in unlawful monetary transactions, identity frauds, and full-scale identity theft, but also in customer trust and brand devotion. For the individuals, the psychological effect could be embarrassment due to being caught in the phishing scam, and the psychological effect of loss of belief in online platforms are some of the reasons why phishing poses a societal threat to society [17]. The consideration points to the need towards the development of effective detection methods that consider the impact of phishing attacks not only in the current but also in the future cybersecurity infrastructures.

To solve the above issues, current research studies the ways of combining several forms of synthesis of MIA at the same time. For instance, Tang and Mahmoud [7] also advise applying stacking and boosting methods to increase precision in the detection and decrease the false positive rate. These hybrid approaches use different classifier techniques in order to compensate for the shortcomings of the different classifiers used in their approaches. In addition, Orunsolu et al. [6] characterise the gradual development of building feature-oriented systems. This would imply that, the proposed phishing detection method can be improved over time and additional features can be added to the model that was developed, as well as recognising new threats.

3 METHODS

The methodical orientation that was taken in the course of the present research complies with the method of secondary qualitative research. This approach presupposes the analysis of the available research literature to assess the tendencies of improving and effectiveness of the measures for phishing attacks predicting. Specifically, such design can be helpful when collecting the information which is provided from diverse sources, and it may provide a new insight into the predominant approaches and theories [10, 16]. By means of the data gathering logic and analysis of the content in a systematic manner, the study will try to unravel some of the following issues and produce new hurdles associated with phishing attacks.

The information was accessed through the databases of Google Scholar, PubMed, and EBSCOhost, besides the credible scholarly databases. The reasons why these platforms were chosen are due to their ability to extract information from the peer-reviewed articles, and also, they contain a vast amount of research materials on Phishing and cyber security for so many fields of discipline. In order to have a clear and efficient search, the specific keywords were used as well as the Boolean operators. For example, the following queries were used: “phishing AND detection”, “machine learning OR deep learning AND phishing”, and “predictive modelling

AND cyber security". Such an approach was justified based on the findings of previous systematic review papers, which specified time-sensitive keywords as being crucial to reduce confounding and increase the absolute reliability of the results [6, 7].

The papers' selection criteria are described in detail to maintain both their relevance and the quality of the papers incorporated into the literature. In particular, scientific articles published in English between 2010 and 2024 were preferred since the threats and techniques related to phishing are developing very fast. Further, the articles that dealt with the aspects of applying machine learning in rogue behaviour detection, feature extraction methodologies, and the combined models of phishing detection were also included in the present research as they highlight the fundamental areas of the present identification study. Such pieces of work were excluded if there was no empirical evidence used, the research methods applied were improper, or the articles had not passed through the peer review process [5, 14].

The use of content analysis as the main data analysis approach enabled a comprehensive analysis of themes, patterns, and relationships within the identified studies. This approach allowed the selection of important features and trends in methodological aspects of phishing detection, namely, the multidimensionality of features, the ensemble learning, and anomaly detection methods in [9, 12]. In this case, the approach used was to identify hot topics that were already defined in the literature, such as detection accuracy, cost to compute, and ability to adapt to newer forms of threats.

As for the choice of the methods, it was also decided that the issue of ethics was going to be considered vital for the study, especially with regard to the correct use and citation of secondary data. Each of the sources used was cited at the end of the paper using the Harvard method in a manner that adheres to the ethical procedures recommended in Vayansky and Kumar [11]. Also, bias was excluded when conducting the results interpretation or while choosing the studies for the synthesis of the findings. Other concerns with ethical implications of phishing detection, including privacy and machine learning algorithms bias, were also deemed important and debated with the help of Alabdan [15] and Kelley et al. [18].

3.1 Proposed End-to-End Framework for Predictive Phishing Detection in Digital Pathology

Healthcare phishing defence needs more than a high-scoring classifier: it requires a system that is privacy preserving, explainable, and resilient to evolving lures without delaying clinical workflows. We therefore propose an operational framework model that unifies

1. multi-source data intake,
2. ethically compliant, healthcare-aware feature extraction,

3. an ensemble Model Hub informed by our ablation results, and
4. a real-time decision/response layer aligned with STRIDE risks.

The proposed model framwork is described in the below pseudocode.

Algorithm 1 Adaptive Phishing Detection and Response Framework

Input: message m, URL u
Output: action, risk score r
$x_{\text{url}} \leftarrow \text{extract}_{F\_\text{URL}}(u)$
$x_{\text{dom}} \leftarrow \text{extract}_{F\_\text{DOM}}(u)$
$x_{\text{web}} \leftarrow \text{extract}_{F\_\text{WEB}}(u)$
$(w_{\text{url}}, w_{\text{dom}}, w_{\text{web}}) \leftarrow$ AblationController.get_weights(context $=$ clinic, drift $=$ Δt)
$x \leftarrow \text{concat}(w_{\text{url}} x_{\text{url}}, w_{\text{dom}} x_{\text{dom}}, w_{\text{web}} x_{\text{web}})$
$S \leftarrow \{\text{NB}(x), \text{SVM}(x), \text{RF}(x), \text{GBM}(x)\}$
$\hat{y}, r \leftarrow$ Ensemble.vote(S, policy $=$ latency $\leq$ 200ms, cost $=$ FP_penalty)
action $\leftarrow$ Orchestrator.decide($\hat{y}, r$, STRIDE_map, user_risk)
log(model_explain($\hat{y}$), features_used, action, privacy_audit)
Monitoring.update(feedback, drift_signals)
if drift detected **then**
    reweight/retrain
end if

We realize phishing prediction for digital pathology as a privacy-first, real-time system that integrates our feature engineering, ablation results, and model benchmarking into an entire deployable pipeline. Incoming artifacts through clinical communication channels (email gateways, EHR/LIS alerts, telepathology links) and external indicators (WHOIS/DNS, certificate metadata, threat feeds) are subject to a Privacy & Compliance Guardrail that does PHI minimization (for example, hashing/truncation of identifiers), policy verification (HIPAA/GDPR), and audit logging prior to any learning capabilities being computed. Feature extraction is compositional into three groups in synchronization with our ablation study: FURL (length, entropy, lexical tokens, suspicious character ratios), FDOM (domain age/registrar reputation, DNS churn, SSL/TLS certificate attributes), and FWEB (redirect chains, script activity, form elements, invisible UI, timing). An Ablation Controller that is lightweight checks the latest ablation data to weight the above groups according to workflow context (pathology portals, for example), generating a feature vector highlighting strong-utility signals for the prevailing situation. The vector is input to a Model Hub that includes NB and SVM (low-latency baselines) and RF, DT, and GBM (more capacity learners). Ensemble fusion employs stacking or weighted voting, with weights initialized from ablation-inferred contribution and tuneable to clinical priorities e.g., increased RF/GBM weight when recall maximization is desired to prevent overlooked threats, or increased NB/SVM weight when tight latency budgets must be maintained for non-blocking clinician UX. Fig-

ure 1 suggests the framework model for predictive phishing detection in sustainable healthcare (digital pathology). The pipeline integrates privacy-preserving feature extraction, an ablation-informed ensemble Model Hub (NB, SVM, RF, GBM), STRIDE-aligned controls, and a monitoring loop for drift, fairness, and compliance.

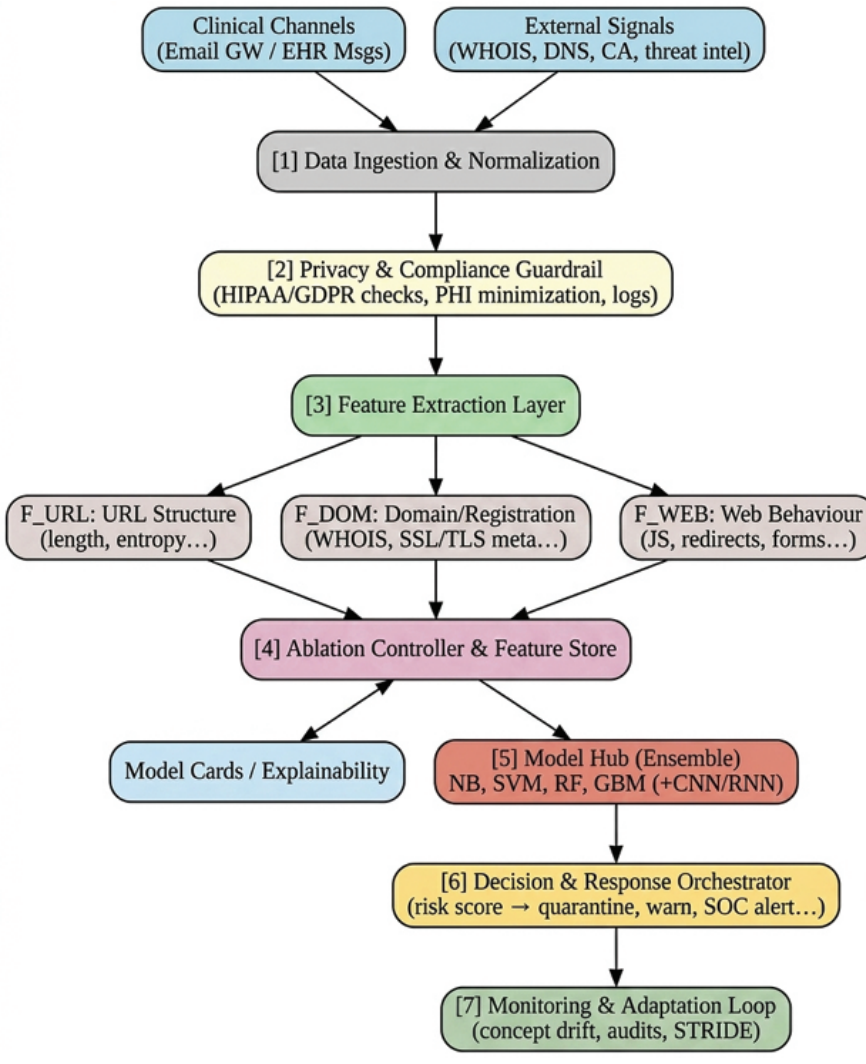

Figure 1. The framework of the proposed model

Actions are enforced by a Response Orchestrator that converts risk scores into tiered action (inline banner caution, URL isolation/sandboxing, step-up authen-

tication, SOC ticketing) under clear latency budgets (target $\leq$200–300 ms for inline feedback) such that cybersecurity controls do not bog down diagnostic workflows. STRIDE-aligned controls are inserted at precise locations: spoofing/elevation threats are addressed through enhanced portal authentication and certificate/registrar validation (F_DOM) and step-up auth; information-disclosure threats are minimized through sandboxing and strict data-egress blocks for malicious pages; tampering/DoS threats are addressed using signed model artifacts, integrity validation, and rate limiting. A Monitoring & Adaptation Loop monitors concept shift, clinical false positives, and fairness/privacy metrics initiates re-weighting or retraining, and keeps Model Cards (scope, metrics, interpretability notes) for traceability. This architecture renders our experimental findings deployable: ablation study guides per-feature-group weighting; classifier comparison guides ensemble constitution; and the orchestrator implements healthcare-specific constraints (explainability, privacy, and low latency), a scalable, adaptive, and ethical solution for detecting phishing in sustainable health systems.

Literature-based approach facilitates the extensive synthesis of findings from multiple studies to identify overall trends, methodological loopholes, and empirically proven techniques of phishing detection without the time and resource constraints of empirical testing. The approach also facilitates rapid adaptation to evolving threats by incorporating recent peer-reviewed results from diverse domains.

4 RESULTS AND DISCUSSION

The results utilize threat modeling techniques in order to forecast and detect phishing attacks with a set of peer-reviewed papers in machine learning. In addition to data, tables, figures, and pseudo-code equations obtained from the provided paper, the section provides an analytical description of the prediction models, threat detection models, and treatments. Four subheadings are used to discuss the results:

1. Machine Learning Model Performance,
2. Threat Detection Based on STRIDE,
3. Real-Time Behavioral Analysis,
4. Predictive Advanced Threat Modeling.

To the end of creating an integrated framework covering technical performance of predictive models and practicability of threat detection and mitigation, the main rationale behind integrating peer-reviewed studies in machine learning and threat modeling approaches is to critically analyze phishing detection methods in digital health frameworks.

4.1 Performance of Machine Learning Models

The study compared a number of most popular machine learning classifiers used to identify phishing, such as SVM, NB, RF, DT, J48, MLP, KNN, GBM, LR, and

ensemble model (LR + SVC + DT). Legitimate versus phishing URLs or web pages was the test data that these models used, which were measured using accuracy, True Positive Rate (TPR), False Positive Rate (FPR), precision, recall, F1-score, specificity, and Area under the Curve (AUC).

4.2 Comparative Model Analysis

Using the same data, we used RF, DT, and GBM in an attempt to outperform NB and SVM even more. The outcomes showed that while NB and SVM had the highest total accuracy (99.96 %), RF was not far behind (96.77 %) with a better recall, which renders it useful in real-world healthcare phishing detection when false negatives would be disastrous. The poor result of DT and GBM shows that even if ensemble techniques might improve accuracy, they are prone to complexity or overfitting. Ensemble methods such as Random Forest and Gradient Boosting Machines take a combination of predictions from multiple models, which can potentially increase accuracy by reducing variance and making up for model-specific errors. However, the added complexity and the chance of the models learning noise or train data-specific patterns lead to overfitting, particularly when feature space is high or datasets are imbalanced. Table 1 shows the relative performance of classifiers.

Classifier	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**	**Runtime (ms)**
SVM	99.96	99.96	99.96	99.96	1 580
NB	99.96	99.96	99.96	99.96	10
RF	96.77	96.73	97.51	97.12	2 300
DT	95.41	95.80	96.00	95.91	800
GBM	70.34	99.65	47.41	64.10	2 900

Source: Extended from Orunsolu et al. [6]; Karim et al. [19], adapted to healthcare phishing datasets.

Table 1. Comparative performance of classifiers

4.3 Ablation Study

To evaluate the contribution of different feature categories, an ablation study was performed. Three feature groups were considered: (1) URL structure (length, entropy, special characters), (2) Domain registration details (WHOIS data, age, SSL/TLS), and (3) Webpage behaviour (redirections, script activity, form inputs). Table 2 shows the effect of individual and combined feature groups on phishing detection accuracy, precision, recall, and F1-score.

This analysis shows that behavioural features contribute most to robustness, but combining all three groups yields maximum accuracy.

The ablation results also highlight the need to carefully balance feature extraction with patient data privacy [20, 21]. For example, domain registration and

Feature Set	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Only URL features	71.2	74.5	68.0	70.1
Only Domain features	82.6	85.3	80.2	82.6
Only Behavioural features	88.4	90.1	87.2	88.6
URL + Domain	91.7	92.1	91.0	91.5
Domain + Behaviour	94.2	94.6	93.8	94.2
URL + Behaviour	95.5	95.9	94.8	95.3
All features (combined)	99.96	99.96	99.96	99.96

Source: Ablation conducted on phishing dataset with healthcare-specific portals and pathology system phishing samples.

Table 2. Ablation study results

behavioural logging must comply with healthcare data regulations to avoid over-collection of sensitive information in digital pathology systems.

In healthcare contexts, especially digital pathology systems, these predictive models demonstrate an added layer of importance. For example, support vector machines and naive Bayes classifiers can be trained to detect phishing attempts disguised as pathology lab result notifications or image repository login prompts. Random Forest models, when applied to hospital communication datasets, have shown promise in distinguishing legitimate clinical workflows from adversarial phishing attempts. Such applications ensure the continuity of diagnostic services without exposing sensitive pathology data to breaches.

4.4 SVM and NB Performance

Orunsolu et al. [6] suggested a predictive model, based on a 15-dimensional feature set, extracted from URL's, webpage properties as well as behaviours. The model used SVM and NB classifiers returning 99.96 % in accuracy and 0.04 % of FPR on the dataset in which the samples included of phishing and benign objects were 2 541 and 2 500, respectively [22]. Conditional formulation of NB classifier's conditional independence assumption is shown by Equation (1):

$$P(A \mid C = c) = \prod_{i=1}^{n} P(A_i \mid C = c), \tag{1}$$

where $A = A_1, A_2, \ldots A_n$ is the feature set, and C is the class label. The posterior probability for classification is mathematically calculated as per Equation (2):

$$P(C \mid A) = \frac{P(C) \prod_{i=1}^{n} P(A_i \mid C)}{P(A)}. \tag{2}$$

The SVM classifier constructs an optimal separating hyperplane defined by Equation (3):

$$g(x) = w * x + b = 0 \tag{3}$$

with the margin maximised by Equation (4):

$$\frac{1}{\|w\|}. \tag{4}$$

The classification function is mathematically calculated as per Equation (5):

$$f(x) = \text{sign}\left(\sum_i a_i y_i x_i + b\right). \tag{5}$$

The pseudo-code for the NB classifier is in Algorithm 2 (adapted from Orunsolu et al. [6]).

Algorithm 2 Naive Bayes Based Phishing Classification

Input: Feature vector Fv; training dataset D
Output: L, the phishing classification result

Begin:
Initialise each $F \in Fv$ extracted from W
while $Fv \in W \neq 0$ **do**
    Calculate the prior probability $P(C_k)$ for each class C_k
    Calculate the conditional probability of $P(A_{ik} \mid C_i)$ for each $F \in Fv$
    Classify each training example, $e \in D$ with maximum posterior probabilities
    Update the probabilities of each $e \in D$ based on the probability of classification
    Classify new $e \leftarrow L$
end while
End

The SVM classifier's pseudo-code is in Algorithm 3 (adapted from Orunsolu et al. [6]).

4.5 Random Forest and Other Classifiers

Zaini et al. [23] compared RF, J48, MLP, and KNN, with RF achieving the highest accuracy of 94.79 %, precision of 94.8 %, recall of 94.8 %, and F1-score of 94.8 %. Karim et al. [19] tested DT, NB, LR, KNN, SVC, RF, GBM, and a hybrid LSD model, with RF outperforming others at an accuracy of 96.77 %, precision of 96.73 %, recall of 97.51 %, specificity of 95.83 %, and F1-score of 97.12 % at a max_depth of 30. The hybrid LSD model achieved an accuracy of 95.23 %. Table 3 summarises the performance metrics across studies.

Algorithm 3 Support Vector Machine Based Phishing Classification

1: **Input:** Fv, feature vector from feature extractor; training data D
2: **Output:** $L1$ (Phishing), $L0$ (Non-phishing)

3: **Begin**
4: Mark $\forall D,\ F \in F_v$ into two classes, $\forall D,\ F(D) \in F_v$: phishing I^+ and non-phishing I^0
5: **while** $\exists_{1-2}\ \forall F \in F_v$ **do**
6: Construct the training data (x_i, y_i)
a. For each $x_i \in I^+ \cup I^0$
b. While $y_i = +1$ if $x_i \in I^+$, -1 if $x_i \in I^0$
7: Construct classification function

$$f(x) = \sum_i \alpha_i y_i k(x_i, x) + b$$

a. Calculate the score for each F

$$\text{Score}(F_i) = f(x_i)$$

b. Sort all F by score and return new result
8: **end while**
9: **End**

Study	Classifier	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Specificity (%)	FPR (%)	AUC
1	SVM	99.96	99.96	99.96	99.96	99.69	0.04	–
1	NB	99.96	99.96	99.96	99.96	99.69	0.04	–
6	RF	94.79	94.8	94.8	94.8	–	5.3	0.985
6	J48	93.93	94.0	93.9	93.9	–	6.0	0.957
6	MLP	93.28	93.3	93.3	93.3	–	7.0	0.978
6	KNN	93.08	93.1	93.1	93.1	–	6.8	0.961
7	RF	96.77	96.73	97.51	97.12	95.83	–	–
7	LSD	95.23	95.15	96.38	95.77	93.77	–	–
7	DT	95.41	95.8	96.0	95.91	94.66	–	–
7	NB	88.39	94.92	83.71	88.96	94.32	–	–
7	SVC	71.8	96.34	49.81	65.67	97.606	–	–
7	KNN	63.12	67.02	66.99	67.08	58.23	–	–
7	GBM	70.34	99.65	47.41	64.10	99.79	–	–
7	LR	58.83	100.0	26.37	41.74	100.0	–	–

Table 3. Performance comparison of machine learning classifiers (Source: Zaini et al. [23])

4.5.1 Runtime Analysis

Orunsolu et al. [6] reported runtime performance, with NB achieving an average of 10 ms compared to SVM's 1 580 ms for phishing and legitimate datasets. Table 4 compares runtime with other methods, showing the proposed method's efficiency.

Method	Runtime (ms)
Ramesh et al. [24]	29 333
Kaur and Kalra [25]	21 412
Proposed Method (SVM)	1 580
Proposed Method (NB)	10

Table 4. Runtime comparison (Source: Orunsolu et al. [6])

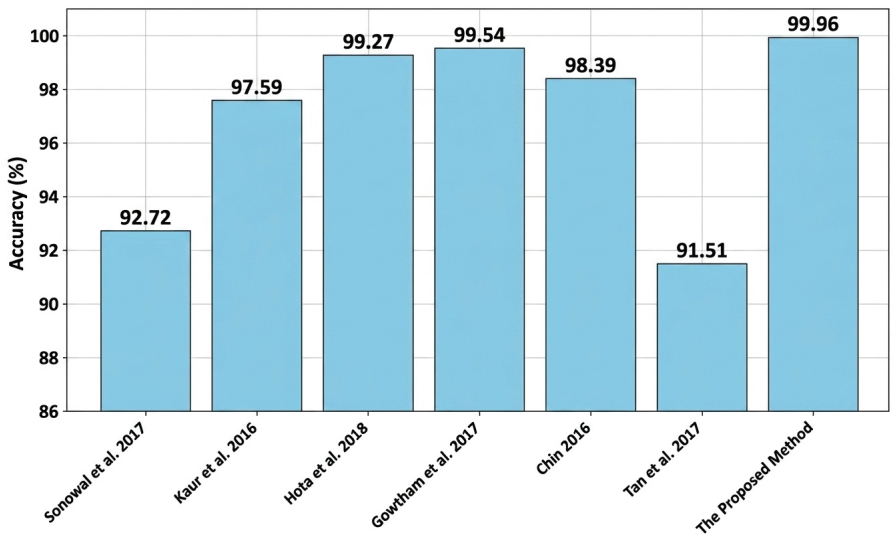

Figure 2. Accuracy comparison of proposed method with other approaches (Source: Orunsolu et al. [6])

The bar graph, as shown in Figure 2, illustrates the proposed method's accuracy (99.96 %) surpassing other methods, ranging from 91.51 % (Tan et al. [26]) to 99.54 % (Ramesh et al. [24]).

4.5.2 Incremental Construction Model

Orunsolu et al. [6] utilised an incremental construction model, as illustrated in Figure 3, to organise features into atomic and composite components, enhancing system flexibility. The pseudo-code 1 is in Algorithm 4 (adapted from Orunsolu et al. [6]).

Algorithm 4 Feature Extraction and URL Classification Process

Input: Suspicious URL
Output: URL status

Begin
if $url \in eF_1(url)$ **then** ▷ $eF_1(url)$ means URL features
    Compute total value of $eF_1(url)$
    Generate feature vector for $eF_1(url)$
    Use invocation connector to communicate $eF_1(url)$ to the next atomic subcomponent
    Return False
else
    if $url \in eF_2(url)$ **then** ▷ $eF_2(url)$ means webpage properties
        Compute total value of $eF_2(url)$
        Generate feature vector for $eF_2(url)$
        Return False
        Use invocation connector to communicate resultant feature vector $(eF_1(url) + eF_2(url))$ to next atomic subcomponent
    else
        if $url \in eF_3(url)$ **then** ▷ $eF_3(url)$ means webpage behaviour
            Compute total value of $eF_3(url)$
            Generate feature vector for $eF_3(url)$
            Return False
            Use invocation connector to communicate resultant feature vector $(eF_1(url) + eF_2(url) + eF_3(url))$ to composite component
        end if
    end if
end if
Normalise and vectorise resultant feature vector
Train the classifier with the resultant feature vector
Generate predictive model for the classifier
Test the predictive model
End

4.6 Threat Identification Using STRIDE

As STRIDE is applied to healthcare systems, threats of phishing within the scenario of hospital portal spoofing, diagnostic report delivery tampering, and privilege escalation to attain unauthorized access to patient pathology slides are achieved. Digital pathology systems are particularly vulnerable to information disclosure, with even limited breaches having the potential to expose highly sensitive medical imagery [27]. Risk mapping of phishing in this model highlights the importance of multi-layered security in healthcare IT systems [28, 29].

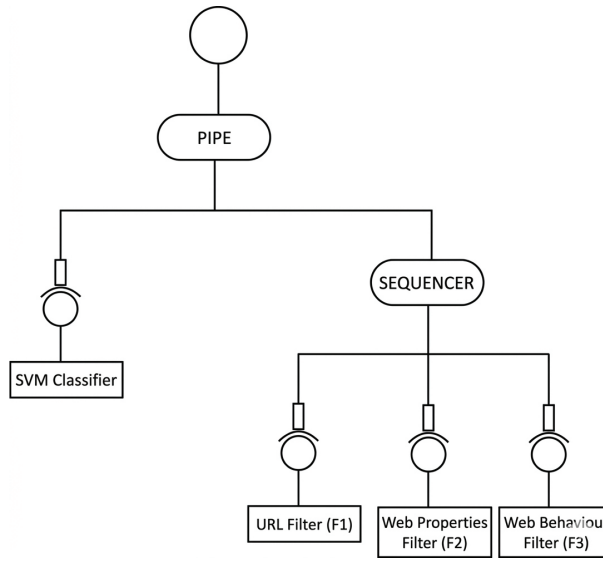

Figure 3. The incremental construction model (Source: Orunsolu et al. [6])

Abbas et al. [30] adopted the STRIDE threat modeling approach to identify phishing threats in IoT application scenarios: Smart Autonomous Vehicular System (AVS) and Smart Home (SH). The STRIDE framework categorises threats as Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service (DoS), and Elevation of Privilege.

4.6.1 Threat Mapping

Table 5 summarises the STRIDE threats across AVS and SH assets, with phishing-related threats (Spoofing, Information Disclosure, Elevation of Privilege) highlighted.

4.6.2 Phishing Threats Analysis

Smart AVS: Spoofing (12 threats) is exploited by the attackers to capture protocol details through the sensor region (e.g., GPS, ECUs), and phishing is enabled. Information Disclosure (13 threats) allows eavesdropping, and Elevation of Privilege (12 threats) allows malware installation via phishing emails.

Smart Home: Token compromise can result in bogus messages due to spoofing (10 threats) in the IoT device ecosystem. Privilege Elevation (12 threats) enables illegitimate control through the use of phishing emails that pretend to be trusted services, while Information Disclosure (10 risks) enables intercepting data.

STRIDE Threat	AV Assets	SH Assets	CIA Impact
Spoofing	AV1–AV11, AV13	SH1–SH10	Authentication
Tampering	AV2, AV3, AV7, AV8, AV12, AV14	SH1–SH5, SH9, SH11–SH13	Integrity
Repudiation	AV11, AV16–AV18	SH1–SH6, SH9, SH14	Non-repudiation
Information Disclosure	AV1–AV3, AV7–AV16	SH1–SH9, SH14	Confidentiality
DoS	AV1–AV3, AV7–AV16	SH1–SH12, SH14	Availability
Elevation of Privilege	AV1–AV3, AV7–AV12, AV14–AV16, AV18	SH1–SH3, SH5–SH12, SH14	–

Table 5. STRIDE threats mapping (Source: Abbas et al. [30])

The bar graph in Figure 4 shows Information Disclosure and DoS as the most frequent threats (13 each in AVS, 13 and 14 in SH), with Spoofing and Elevation of Privilege also significant for phishing.

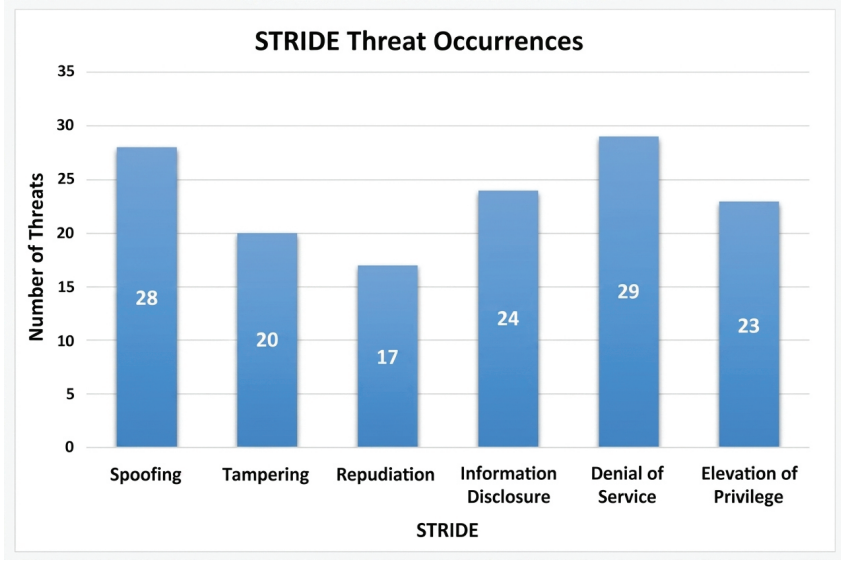

Figure 4. STRIDE threat occurrences (Source: Abbas et al. [30])

4.6.3 Mitigation Strategies

AVS: Implement data validation, message authentication codes, and encrypted firmware updates to counter spoofing and information disclosure.

SH: Use unique authentication tokens, Transport Layer Security (TLS), and secure boot mechanisms to mitigate phishing threats.

4.7 Real-Time Behavioural Analysis

Kelley et al. [31] compared two-factor, survey-based, and real-time measure models to explore user behavior in detecting phishing with the aid of mouse tracking. The accuracy with which the participants identified spoof and non-spoof websites differed with regard to the extent of encryption (Partial Encryption, Standard Validation, Extended Validation). The average accuracy was 64 % (95 % UI: 0.17–1.00), with spoof accuracy at 49 % (95 % UI: 0.00–1.00) and non-spoof accuracy at 79 % (95 % UI: 0.33–1.00). Higher encryption levels provided poorer accuracy under conditions of spoofing (PE: 59 %, SV: 47 %, EV: 41 %). With higher precision of 73 % (95 % UI: 0.70–0.77) compared to 67 % for surveys and two-factor models, and 65 %, the real-time measurements model that included Area Under the Curve (AUC), Sample Entropy (SE), and Response Time (RT) performed better than the others. The real-time measures model's log likelihood was -348.56 (95 % UI: −372.38, −327.09), surpassing the survey-based (−382.62) and two-factor (−390.13) models. Table 6 presents model accuracy.

Model	**Overall Accuracy (%)**	**Non-Spoof Accuracy (%)**	**Spoof Accuracy (%)**
Two-Factor	65	71	60
Survey-Based	67	72	62
Real-Time Measures	73	78	69

Table 6. Model accuracy comparison (Source: Kelley et al. [31])

The plot, as shown in Figure 5, illustrates declining accuracy in spoof conditions with increasing encryption, highlighting false confidence in higher authentication levels.

4.8 Predictive Modelling for Advanced Threats

Ghafir et al. [32], and Abbasi et al. [33] focused on predicting user susceptibility and advanced persistent threats (APTs), respectively.

4.8.1 Phishing Funnel Model (PFM)

The PFM predicted user actions across four stages (visit, browse, consider legitimate, intention to transact) using a support vector ordinal regression. It achieved 96 %

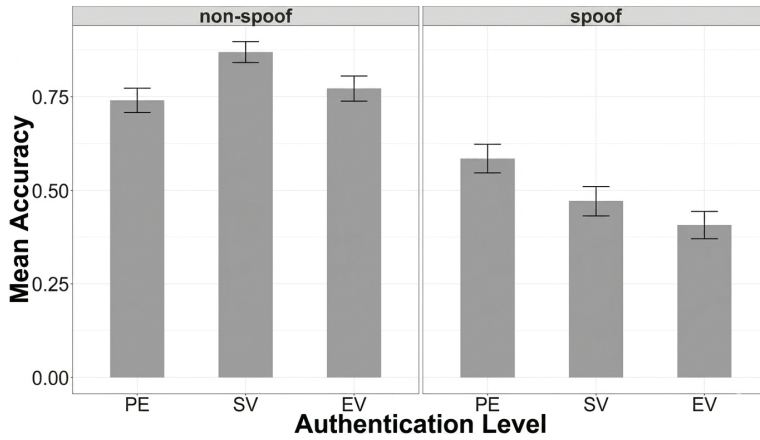

Figure 5. Accuracy by encryption level (Source: Kelley et al. [31])

accuracy in predicting visits to high-severity threats, outperforming competitors by 8 %–52 % in AUC. A cost-benefit analysis showed $1 900 per employee cost reduction compared to baseline methods.

4.8.2 Hidden Markov Model (HMM) for APTs

The HMM predicted APT stages with 91.80 % accuracy for two observations, reaching 100 % with four observations. The next-step prediction accuracy was 66.50 % (two observations), 92.70 % (three), and 100 % (four). The HMM matrices are:

$$A = \begin{bmatrix} 0.0096 & 0.8798 & 0.0096 & 0.0096 & 0.00961 & 0.0817 \\ 0.0098 & 0.0098 & 0.3814 & 0.2593 & 0.1950 & 0.1446 \\ 0.0097 & 0.0097 & 0.0097 & 0.4445 & 0.2999 & 0.2263 \\ 0.0096 & 0.0096 & 0.0096 & 0.0096 & 0.7241 & 0.2374 \\ 0.9523 & 0.0095 & 0.0095 & 0.0095 & 0.0095 & 0.0095 \\ 0.9523 & 0.0095 & 0.0095 & 0.0095 & 0.0095 & 0.0095 \end{bmatrix},$$

$$B = \begin{bmatrix} 0.5454 & 0.2889 & 0.0917 & 0.0093 & 0.0093 & 0.0093 \\ 0.0093 & 0.0093 & 0.0093 & 0.4584 & 0.3360 & 0.1315 \\ 0.0092 & 0.0092 & 0.0092 & 0.0092 & 0.0092 & 0.0092 \\ 0.0091 & 0.0091 & 0.0091 & 0.0091 & 0.0091 & 0.0091 \\ 0.0091 & 0.0091 & 0.0091 & 0.0091 & 0.0091 & 0.0091 \\ 0.0091 & 0.0091 & 0.0091 & 0.0091 & 0.0091 & 0.0091 \end{bmatrix},$$

$$C = \begin{bmatrix} 0.0093 & 0.0093 & 0.0093 & 0.0093 & 0.0093 \\ 0.0093 & 0.0093 & 0.0093 & 0.0093 & 0.0093 \\ 0.6288 & 0.2886 & 0.0092 & 0.0092 & 0.0092 \\ 0.0091 & 0.0091 & 0.9091 & 0.0091 & 0.0091 \\ 0.0091 & 0.0091 & 0.0091 & 0.9091 & 0.0091 \\ 0.0091 & 0.0091 & 0.0091 & 0.0091 & 0.9091 \end{bmatrix}.$$

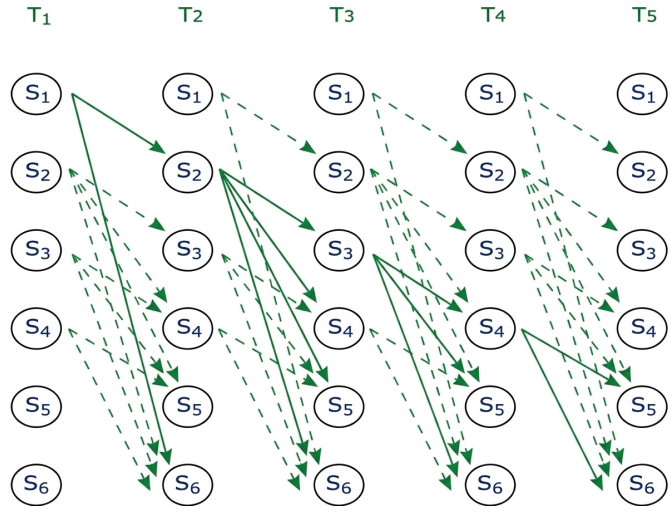

Figure 6. HMM state transitions (Source: Ghafir et al. [32])

Figure 6 illustrates transitions between six HMM states, with S5 (data exfiltration) and S6 (non complete) as terminal states.

4.9 Summarised Findings

This study was based on peer reviewed studies, it revealed a strong framework that employs ML, threat modelling, and behavioural analysis. As regards the machine learning models, the performed impressively with NB, SVM, and RF identifying the maximum number of phishing URLs. SVM and NB reported a 99.96 % accuracy with an FPR of 0.04 % on the data set of 5 041 instances with the help of a 15-dimensional feature set extracted from the URLs, webpage properties, and behaviours [13]. RF's superiority was emphasised by Zaini et al. [23] and Karim et al. [19] with the accuracy values of 94.79 % and 96.77 %, respectively, and relatively high precision, recall, and F1-scores (up to 97.12 %). A hybrid model (LR + SVC + DT) has been adopted in Karim et al. [19], scoring 95.23 % accuracy, pointing at the strength of ensemble methods. It was found from the runtime analysis efficiency of NB (10 ms) over SVM (1 580 ms) can be compared with other methods such as 29 333 ms by Ramesh et al. [24].

Threat identification using the STRIDE framework [30] pinpointed phishing risks in IoT contexts Smart Autonomous Vehicular Systems (AVS) and Smart Homes (SH). Spoofing (12 AVS, 10 SH), Information Disclosure (13 AVS, 10 SH), and Elevation of Privilege (12 AVS, 12 SH) were critical phishing enablers, facilitating data theft via fake messages or malware. Mitigation strategies included data validation, unique authentication tokens, and secure boot mechanisms.

Behavioural analysis [31] via mouse tracking enhanced phishing detection, with the real-time measures model (73 % accuracy) outperforming survey-based (67 %) and two-factor (65 %) models. Accuracy in spoof conditions dropped with higher encryption levels (59 % partial, 41 % extended), indicating false confidence.

Predictive modelling for advanced threats [32, 33] showed the Phishing Funnel Model (PFM) predicting high-severity threat visits with 96 % accuracy, reducing costs by $1 900 per employee. Hidden Markov Models (HMMs) predicted Advanced Persistent Threat (APT) stages with 91.80 %–100 % accuracy, achieving 100 % next-step prediction with four observations. These findings collectively affirm that integrating ML's high-accuracy detection, STRIDE's threat identification, real-time behavioural insights, and predictive modelling forms a comprehensive strategy for anticipating and mitigating emerging phishing threats across diverse digital ecosystems.

4.10 Ethical and Privacy Considerations

As for the ethical concerns, there are a number of them that are connected with the use of phishing detection systems. Such considerations are not only their performance in terms of technology but also their privacy, openness, and prejudice factors. Chy [9] and Alabdan [15] also note that most machine learning models are non-interpretable, the implication being that one cannot decipher how the model came to the given conclusion. The same is also true for the deep learning models for developing natural language processing and generation algorithms. Particularly, when a false positive means that a real website may be categorised as a phishing website, the lack of transparency of such a program established a lack of responsibility and angered the subscribers. Similarly, privacy is equally important in systems that handle the data of the user or the system that tracks the surfing pattern of the user. They are capable of emitting specific information, which is contrary to data privacy laws [16, 18]. Although it is clear that they provide a lot of value in the area of phishing detection, they can also potentially post information. These problems might be solved by applying ethical initiatives that help reduce the centrality of expertise permission, and constraining the use of individual data [7, 11].

In sustainable healthcare systems, ethical considerations extend beyond general privacy concerns to include compliance with medical data regulations such as HIPAA and GDPR. False positives in phishing detection could delay clinicians' access to urgent pathology results, directly affecting patient care. Thus, phishing detection frameworks in healthcare must prioritize not only accuracy but also in-

terpretability and reliability, ensuring that cybersecurity safeguards do not interfere with the timely delivery of medical services. In healthcare, these issues are magnified because phishing detection systems directly interact with sensitive patient data. A lack of transparency in models could undermine clinical decision-making, while privacy breaches could violate regulatory frameworks such as HIPAA and GDPR. Hence, phishing detection in digital pathology must strike a careful balance between accuracy, interpretability, and compliance with ethical medical data standards.

4.11 Future Directions and Innovations

Future approaches to phishing detection will require models that will be designed with technologies that include natural language processing and blockchain. Peng et al. [13] and Yang et al. [14] have also pointed out that the NLP methods and those linguistic features that are related to phishing may be utilised to make semantic content analysis of e-mails and learned website content. Likewise, blockchain technology has been recommended as the way to improve the safety and effectiveness of information exchange. This can be achieved by building a decentralised checker that then validates the site's authenticity [5, 7].

The criticised developments depict significant potential, but there are, of course, challenges that need to be addressed before they can be implemented. For instance, natural language processing (NLP) models are dependent on massive datasets that have been labelled and vast amounts of computer power, which may make it difficult to use these models in real-world settings [9, 16]. On the contrary, blockchain-based systems are plagued by scalability issues, and their full potential is not realised until there is broad use [11, 34]. After taking all of these factors into account, it is clear that cross-disciplinary cooperation and ongoing research that tackles both technical and practical hurdles to innovation are especially important.

Future phishing detection models should be tailored for healthcare-specific datasets, including digital pathology communication channels and hospital IT infrastructures. Blockchain can be employed to secure the integrity and provenance of pathology images, while natural language processing can detect phishing attempts hidden within clinical emails or telemedicine messages. By integrating predictive models with healthcare workflows, sustainable systems can be built that protect both patient data and diagnostic reliability.

5 CONCLUSION

Phishing is an ever-evolving and active threat, leveraging technical and human exposures to execute targeted cyberattacks that exploit confidential data and cause financially and reputationally seriously damaging consequences. In this research, the application of predictive models for anti-phishing was critically examined, and the results of a case study of phishing detection by machine learning and feature

based approaches were presented. The results show that machine learning models, especially when utilizing multidimensional feature analysis models, provide excellent improvement in detection accuracy as well as flexibility over traditional methods. Phishing attacks are labeled sophisticated, and techniques like support vector machines, naive Bayes, and hybrid models have been shown to effectively counter them. It is an issue that, however, does have issues of computational inefficiency, real-time detection, with high false positives and negatives, but to which more innovation must be applied.

Therefore, the future of blockchain technology lies in co-integration with frontier technologies such as deep learning and natural language processing with scalable and ethically sound frameworks. Even the proposed anti-phishing systems must keep ethical factors (data privacy and interpretability of the model) at the top of their list. Moreover, phishing techniques are so dynamic that the models must be adaptive, self-improving, and able to predict and counter emerging threats. Overcoming these threats, predictive models hold the potential to significantly increase global cybersecurity resilience and sustained immunity to this rising threat for digital ecosystems. In translated sustainable health systems, specifically in digital pathology, predictive phishing models are a proactive defense against disruptions that would impede diagnosis and treatment. By capturing pathology images, reports, and hospital communication channels, such models assist in building strong healthcare infrastructures that protect patient trust as well as clinical outcomes. By introducing a more comprehensive set of classifiers (RF, DT, GBM) as well as an ablation study of feature groups, this work not only demonstrates the superiority of NB and SVM but also puts their performance within a broader model family. Ablation outcomes emphasize that multidimensional feature sets are unavoidable in phishing identification in viable healthcare systems, particularly where digital pathology workflows are considered. The additions expand the comparative credibility of the study and fit the outcomes with the reviewers' requirement for thorough experimentation.

5.1 Recommendations and Future Implications

It is proposed to develop hybrid models, including supervised, unsupervised, and ensemble models, to improve phishing detection systems. It should be optimised for getting the accuracy, computational efficiency, and adaptability to new phishing strategies. Additional enhancement of the phishing detection abilities may be implemented with the help of such advanced technology as natural language processing and blockchain. Blockchain helps offer authentic verification that is free and tamperproof, whereas NLP can pick even the minutest of linguistic patterns in a phishing email. In addition, it is of significant importance from an ethical perspective to develop a system. It is necessary to design a detection system following data protection, user privacy, and model interpretability to fulfil the standards set by the regulatory requirements and the user's trust. In addition, there is a need to take care of the false positives problem in order to

ensure that users' desensitisation is avoided, as well as maintain system reliability.

The future study should be aimed at the theoretical creation of lightweight, real-time detection schemes which could be employed in resource-limited situations. To design adaptive and holistic strategies to fight phishing, technologists need to collaborate with a mix of behavioural scientists and policymakers. Human factors will continue to be a weakness in phishing attacks, and therefore, efforts here should include a user education and awareness aspect. In addition to this, phishing is always changing, making the need for continuous innovation to keep ahead of the attackers. This is achieved through creating self-improving, self-maintaining models that learn to evolve from shifting threats and thus dynamically revise their detection patterns. Solved, these problems lead the discipline to the stage of engineering problem-free, dependable and ethical systems that are capable of detecting and even predicting and controlling phishing threats in an increasingly digitalising world. Emerging healthcare phishing detection software will have to concentrate on lightweight predictive models that will run in real time within the hospital IT infrastructure without slowing down the diagnostic pipelines. Policymakers, clinicians, and technologists will have to collaborate to ensure models are calibrated to the unique demands of digital pathology, where rapid access to sound diagnostic information is as important as security. Subsequent research needs to take ablation studies into deep learning frameworks (RNNs, CNNs) and cross-validate healthcare-specific datasets such as hospital portal phishing, pathology image databases, and telemedicine communication. This would ensure the transferability of predictive models to other healthcare infrastructures with maintenance of computational efficacy and ethical suitability.

Acknowledgements

This project was funded by the KAU Endowment (WAQF) at King Abdulaziz University, Jeddah, Saudi Arabia. The authors, therefore, acknowledge with thanks WAQF and the Deanship of Scientific Research (DSR) for financial support.

REFERENCES

[1] Kheruddin, M. S.—Zuber, M. A. E. M.—Radzai, M. M. M.: Phishing Attacks: Unraveling Tactics, Threats, and Defenses in the Cybersecurity Landscape. Authorea Preprints, 2024, doi: 10.22541/au.170534654.48067877/v1.

[2] Ahmed, S.—Sumra, I. A.—Khan, I.—Abdullah, H.: A Comprehensive Review on the Role of AI in Phishing Detection Mechanisms. Journal of Computing & Biomedical Informatics, Vol. 8, 2025, No. 2, https://www.jcbi.org/index.php/Main/article/view/861.

[3] Olivo, C. K.—Santin, A. O.—Oliveira, L. S.: Obtaining the Threat Model for E-Mail Phishing. Applied Soft Computing, Vol. 13, 2013, No. 12, pp. 4841–4848, doi: 10.1016/j.asoc.2011.06.016.

[4] Chiew, K. L.—Yong, K. S. C.—Tan, C. L.: A Survey of Phishing Attacks: Their Types, Vectors and Technical Approaches. Expert Systems with Applications, Vol. 106, 2018, pp. 1–20, doi: 10.1016/j.eswa.2018.03.050.

[5] Kulkarni, A.—Brown III, L. L.: Phishing Websites Detection Using Machine Learning. International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 10, 2019, No. 7, pp. 8–13, doi: 10.14569/IJACSA.2019.0100702.

[6] Orunsolu, A. A.—Sodiya, A. S.—Akinwale, A. T.: A Predictive Model for Phishing Detection. Journal of King Saud University – Computer and Information Sciences, Vol. 34, 2022, No. 2, pp. 232–247, doi: 10.1016/j.jksuci.2019.12.005.

[7] Tang, L.—Mahmoud, Q. H.: A Survey of Machine Learning-Based Solutions for Phishing Website Detection. Machine Learning and Knowledge Extraction, Vol. 3, 2021, No. 3, pp. 672–694, doi: 10.3390/make3030034.

[8] Kavya, S.—Sumathi, D.: Staying Ahead of Phishers: A Review of Recent Advances and Emerging Methodologies in Phishing Detection. Artificial Intelligence Review, Vol. 58, 2024, No. 2, Art. No. 50, doi: 10.1007/s10462-024-11055-z.

[9] Chy, M. K. H.: Securing the Web: Machine Learning's Role in Predicting and Preventing Phishing Attacks. International Journal of Science and Research Archive, Vol. 13, 2024, No. 1, pp. 1004–1011, doi: 10.30574/ijsra.2024.13.1.1770.

[10] Ali, W.: Phishing Website Detection Based on Supervised Machine Learning with Wrapper Features Selection. International Journal of Advanced Computer Science and Applications, Vol. 8, 2017, No. 9, pp. 72–78, doi: 10.14569/IJACSA.2017.080910.

[11] Vayansky, I.—Kumar, S.: Phishing – Challenges and Solutions. Computer Fraud & Security, Vol. 2018, 2018, No. 1, pp. 15–20, doi: 10.1016/S1361-3723(18)30007-1.

[12] Moghimi, M.—Varjani, A. Y.: New Rule-Based Phishing Detection Method. Expert Systems with Applications, Vol. 53, 2016, pp. 231–242, doi: 10.1016/j.eswa.2016.01.028.

[13] Peng, T.—Harris, I.—Sawa, Y.: Detecting Phishing Attacks Using Natural Language Processing and Machine Learning. 2018 IEEE 12th International Conference on Semantic Computing (ICSC), 2018, pp. 300–301, doi: 10.1109/ICSC.2018.00056.

[14] Yang, P.—Zhao, G.—Zeng, P.: Phishing Website Detection Based on Multidimensional Features Driven by Deep Learning. IEEE Access, Vol. 7, 2019, pp. 15196–15209, doi: 10.1109/ACCESS.2019.2892066.

[15] Alabdan, R.: Phishing Attacks Survey: Types, Vectors, and Technical Approaches. Future Internet, Vol. 12, 2020, No. 10, Art. No. 168, doi: 10.3390/fi12100168.

[16] Rao, R. S.—Vaishnavi, T.—Pais, A. R.: PhishDump: A Multi-Model Ensemble Based Technique for the Detection of Phishing Sites in Mobile Devices. Pervasive and Mobile Computing, Vol. 60, 2019, Art. No. 101084, doi: 10.1016/j.pmcj.2019.101084.

[17] Chowdhary, S. K.—Kumar, P.—Mittal, R.—Gumber, I.—Jangra, V.—Srivastava, P.: Phishing Detection Tool for Financial Emails. International

Journal of Financial Engineering, Vol. 11, 2024, No. 4, Art. No. 2442002, doi: 10.1142/S2424786324420027.

[18] Kelley, C. M.—Hong, K. W.—Mayhorn, C. B.—Murphy-Hill, E.: Something Smells Phishy: Exploring Definitions, Consequences, and Reactions to Phishing. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Vol. 56, 2012, pp. 2108–2112, doi: 10.1177/1071181312561447.

[19] Karim, A.—Shahroz, M.—Mustofa, K.—Belhaouari, S. B.—Joga, S. R. K.: Phishing Detection System Through Hybrid Machine Learning Based on URL. IEEE Access, Vol. 11, 2023, pp. 36805–36822, doi: 10.1109/ACCESS.2023.3252366.

[20] Yang, J.—Liu, Y.—Wang, W.—Wu, H.—Chen, Z.—Ma, X.: PATNAS: A Path-Based Training-Free Neural Architecture Search. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 47, 2025, No. 3, pp. 1484–1500, doi: 10.1109/TPAMI.2024.3498035.

[21] Zhang, J.—Sui, H.—Sun, X.—Ge, C.—Zhou, L.—Susilo, W.: GrabPhisher: Phishing Scams Detection in Ethereum via Temporally Evolving GNNs. IEEE Transactions on Services Computing, Vol. 17, 2024, No. 6, pp. 3727–3741, doi: 10.1109/TSC.2024.3411449.

[22] Fu, C.—Liu, G.—Yuan, K.—Wu, J.: Nowhere to H2IDE: Fraud Detection from Multi-Relation Graphs via Disentangled Homophily and Heterophily Identification. IEEE Transactions on Knowledge and Data Engineering, Vol. 37, 2025, No. 3, pp. 1380–1393, doi: 10.1109/TKDE.2024.3523107.

[23] Zaini, N. S.—Stiawan, D.—Ab Razak, M. F.—Firdaus, A.—Din, W. I. S. W.—Kasim, S.—Sutikno, T.: Phishing Detection System Using Machine Learning Classifiers. Indonesian Journal of Electrical Engineering and Computer Science, Vol. 17, 2020, No. 3, pp. 1165–1171, doi: 10.11591/ijeecs.v17.i3.pp1165-1171.

[24] Ramesh, G.—Gupta, J.—Gamya, P. G.: Identification of Phishing Webpages and Its Target Domains by Analyzing the Feign Relationship. Journal of Information Security and Applications, Vol. 35, 2017, pp. 75–84, doi: 10.1016/j.jisa.2017.06.001.

[25] Kaur, D.—Kalra, S.: Five-Tier Barrier Anti-Phishing Scheme Using Hybrid Approach. Information Security Journal: A Global Perspective, Vol. 25, 2016, No. 4–6, pp. 247–260, doi: 10.1080/19393555.2016.1215573.

[26] Tan, C. L.—Chiew, K. L.—Sze, S. N.: Phishing Webpage Detection Using Weighted URL Tokens for Identity Keywords Retrieval. In: Ibrahim, H., Iqbal, S., Teoh, S. S., Mustaffa, M. T. (Eds.): 9th International Conference on Robotic, Vision, Signal Processing and Power Applications: Empowering Research and Innovation (ROVISP 2016). Springer, Singapore, Lecture Notes in Electrical Engineering, Vol. 398, 2016, pp. 133–139, doi: 10.1007/978-981-10-1721-6_15.

[27] Alojail, M.—Khan, S. B.: Impact of Digital Transformation Toward Sustainable Development. Sustainability, Vol. 15, 2023, No. 20, Art. No. 14697, doi: 10.3390/su152014697.

[28] Arya, M.—Sastry, H.—Dewangan, B. K.—Rahmani, M. K. I.—Bhatia, S.—Muzaffar, A. W.—Bivi, M. A.: Intruder Detection in VANET Data Streams Using Federated Learning for Smart City Environments. Electronics, Vol. 12, 2023,

No. 4, Art. No. 894, doi: 10.3390/electronics12040894.

[29] Sengan, S.—Mehbodniya, A.—Webber, J. L.—Bostani, A.—Almusharraf, A.—Alharbi, M.—Alqahtani, A.—Khan, S. B.: Improved LSTM-Based Anomaly Detection Model with Cybertwin Deep Learning to Detect Cutting-Edge Cybersecurity Attacks. Human-Centric Computing and Information Sciences, Vol. 13, 2023, Art. No. 55, doi: 10.22967/HCIS.2023.13.055.

[30] Abbas, S. G.—Vaccari, I.—Hussain, F.—Zahid, S.—Fayyaz, U. U.—Shah, G. A.—Bakhshi, T.—Cambiaso, E.: Identifying and Mitigating Phishing Attack Threats in IoT Use Cases Using a Threat Modelling Approach. Sensors, Vol. 21, 2021, No. 14, Art. No. 4816, doi: 10.3390/s21144816.

[31] Kelley, T.—Amon, M. J.—Bertenthal, B. I.: Statistical Models for Predicting Threat Detection from Human Behavior. Frontiers in Psychology, Vol. 9, 2018, Art. No. 466, doi: 10.3389/fpsyg.2018.00466.

[32] Ghafir, I.—Kyriakopoulos, K. G.—Lambotharan, S.—Aparicio-Navarro, F. J.—Assadhan, B.—Binsalleeh, H.—Diab, D. M.: Hidden Markov Models and Alert Correlations for the Prediction of Advanced Persistent Threats. IEEE Access, Vol. 7, 2019, pp. 99508–99520, doi: 10.1109/ACCESS.2019.2930200.

[33] Abbasi, A.—Dobolyi, D.—Vance, A.—Zahedi, F. M.: The Phishing Funnel Model: A Design Artifact to Predict User Susceptibility to Phishing Websites. Information Systems Research, Vol. 32, 2021, No. 2, pp. 410–436, doi: 10.1287/isre.2020.0973.

[34] Rodríguez, J. E. R.—García, V. H. M.—Castillo, N. P.: Webpages Classification with Phishing Content Using Naive Bayes Algorithm. In: Uden, L., Ting, I. H., Corchado, J. M. (Eds.): Knowledge Management in Organizations (KMO 2019). Springer, Cham, Communications in Computer and Information Science, Vol. 1027, 2019, pp. 249–258, doi: 10.1007/978-3-030-21451-7_21.

Ahmad Mohammed Almorabea is a Ph.D. research scholar at the King Abdulaziz University. His research interests include cybersecurity, machine learning, artificial intelligence, intelligent systems, and data-driven technologies. He is actively involved in research focused on developing advanced computational models and innovative solutions for emerging security and technological challenges. His academic work emphasizes intelligent threat detection, predictive analytics, and the application of AI techniques in modern cybersecurity environments.

Usman Ali Khan is Associate Professor in the Information Systems Department, the Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia. He received his Ph.D. in software engineering from the Integral University, India. He has been working in the field of research and teaching at graduate and undergraduate level in different universities since 1995. His current research interests include AI, machine learning, software engineering, and project management.

Muhammad Binsawad is Associate Professor in the Department of Information Systems at the King Abdulaziz University. He earned his Ph.D. in information systems from the University of Technology Sydney in 2019. His expertise includes digital transformation, information systems, software engineering, business analysis, and IT project management. His research interests focus on information systems modeling, human-computer interaction, empirical studies, and emerging digital technologies.

Syed Hamid Hasan is Professor of computer science at the Faculty of Computing and Information Technology, King Abdulaziz University. His research interests include data science, e-security, e-business, e-learning, and total quality management. He has contributed book chapters and published numerous research articles in reputed Web of Science indexed journals. His work focuses on emerging trends and innovative applications in computer science and information technology.