iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** III **Month of publication:** March 2026

DOI: https://doi.org/10.22214/ijraset.2026.78706

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

An Advanced Meta-Learning Ensemble Framework for Interpretable Phishing Detection via RSTHFS Optimization

Dr. Rahul M. Dhokane, Mr. Wakchaure Sanchit Sanjay, Miss. Kale Jayshree Sandip, Miss. Dange Shreya Rajesh, Mr. Wakchaure Ganesh Shivaji
Department of Information Technology Sir Visvesvaraya Institute of Technology, Nashik

Abstract: ***Phishing remains a dominant and highly sophis- ticated form of cybercrime, where attackers deploy deceptive websitestotrickusersintorevealingsensitiveinformation, such as passwords and financial credentials [1], [5]. Despite significant advancements in cybersecurity, accurately detecting these malicious domains remains a critical challenge due to the lack of universally accepted identification parameters and the rapid emergence of ”zero-day” phishing sites [1], [6]. This paper introduces an advanced detection framework that integrates Rough Set Theory-based Hybrid Feature Selection (RSTHFS) withanInnovativeMeta-Learning-BasedEnsembleapproach[3], [6].***

The proposed methodology utilizes a multi-layer stacking architecturetocapturebothglobalnon-linearandlocalpatterns, leveraging base learners such as Residual Multi-Layer Percep- trons (ResMLP) and XGBoost, which are aggregated by a meta-classifier to enhance predictive stability [1], [3]. To ensure the system is lightweight enough for real-time browser deployment, the RSTHFS method is employed to identify a ”minimal reduct” of features, successfully reducing the computational featurespace by over 60% while maintaining high reliability [5], [6]. Furthermore, the framework incorporates Explainable AI (XAI) through SHAP values to provide granular transparency into the model’s decision-making process [6]. Experimental evaluations on benchmark datasets demonstrate a peak accuracy of 98.4%, providing a scalable, efficient, and interpretable solution for modern web security [3], [5].

Index Terms: Phishing Detection, Meta-Learning, Ensemble Learning, SHAP, Explainable AI (XAI), RSTHFS, Browser Security, Stacking Classifier.

I. INTRODUCTION

Phishing represents one of the most pervasive and finan- ciallydamagingformsofcybercrime,leveragingsophisticated social engineering and technical deception to steal sensitive information such as usernames, passwords, and credit card details [1], [5]. Unlike traditional hacking, phishing exploits the ”weakest link” in the security chain—the human user—by creating a sense of urgency or fear through deceptive emails and website pop-ups [3], [4]. As the digital economy grows, the volume of these attacks has reached unprecedented levels, with over 1.35 million new phishing sites appearing globally in recent years [1], [2].

Historically,thedefenseagainstphishingreliedonblacklist- basedsystems,suchasGoogleSafeBrowsing,whichmaintain a database of known malicious URLs [2], [4]. However, these systems are fundamentally reactive and fail to address ”zero- hour” attacks, where a malicious site is active for only a few hours before being taken down and moved to a new domain [2], [3]. This limitation has necessitated the shift toward Machine Learning (ML) and Deep Learning (DL), which can proactively analyze a website’s intrinsic features—such as its URL structure and SSL certificate status—to identify fraud without prior knowledge of the domain [3], [5].

Despite the high accuracy of ML models, two major challenges remain in real-world deployment: computational latency and model interpretability [5], [6]. High-performance deep learning models often require significant processing power, making them difficult to integrate into lightweight browser environments like Chrome extensions [5]. Further- more, most models act as ”black boxes,” providing no ex- planation for why a site was blocked, which can lead to user frustration or ignored warnings [6].

This research addresses these gaps by proposing an ad- vanced framework that integrates Rough Set Theory-based HybridFeatureSelection(RSTHFS)withanInnovativeMeta- Learning Ensemble [3], [5]. By building upon our previous hybrid detection models [6], we introduce a system that optimizesfeaturesetsforspeedwhileutilizingExplainableAI (XAI) to provide transparent justifications for every security alert [6]. This combination ensures a defense mechanism that is not only highly accurate (98.4%) but also fast enough for real-timebrowserfilteringandtransparentenoughforend-user trust [3], [6].

A. Key Terms Defined

- Phishing: An attempt to obtain confidential information by impersonating a credible webpage [5].
- Meta-Learning:Ahierarchicalensembletechniquewhere a meta-learner aggregates predictions from multiple base models to improve overall generalization [3].
- RSTHFS: A feature selection method based on RoughSet Theory that identifies the minimal reduct [5].
- ExplainableAI(XAI):MethodssuchasSHAPthatallow users to understand model outputs [6].

Fig.1.LifecycleofPhishingAttack

II. LITERATURE SURVEY

The evolution of phishing detection has undergone a signif- icant transformation over the past decade. Initially, detection techniquesreliedonsimpleheuristic-basedfilteringandblack- list mechanisms. However, with the increasing sophistication of phishing attacks, these traditional approaches proved in- sufficient. As a result, modern research has shifted towards advanced Machine Learning (ML) and Deep Learning (DL) techniquesthatcanautomaticallylearnpatternsandgeneralize across unseen phishing attempts. This section reviews key methodologies proposed in recent high-impact studies, form- ingthefoundationfortheproposedmeta-learningframework.

A. Meta-LearningandEnsembleInnovation

Recent studies highlight that relying on a single classifier often leads to suboptimal performance due to the highly diverse and evolving nature of phishing datasets. Phishing URLs vary significantly in structure, intent, and obfuscation techniques,makingitdifficultforasinglemodeltogeneralize effectively.

Naseeb et al. (2025) proposed an innovative meta-learning ensembleframeworkthatintegratesArtificialNeuralNetworks (ANN) with Bagging-based K-Nearest Neighbors (KNN), us- ingLogisticRegressionasameta-classifier[3].Thisapproach follows a “learning-to-learn” paradigm, where multiple base learners capture different aspects of the data. ANN effectively models global non-linear relationships, while KNN captures local neighborhood patterns. The meta-classifier then intelli- gently combines these outputs, resulting in improved general- ization and robustness. Their model achieved an accuracy of 97.20%.

Similarly, Kalabarige et al. (2023) introduced a multi-layer stackedensemblemodelenhancedwithboosting-basedhybrid feature selection techniques [1]. Their research emphasized that ensemble models outperform standalone classifiers such as Random Forest and Support Vector Machines (SVM), es- pecially when dealing with imbalanced datasets. The stacking mechanism allows multiple models to collaborate, reducing variance and bias simultaneously, thereby improving predic- tion stability.

B. DeepLearningandSequenceAnalysis

With the increasing complexity of URL obfuscation, deep learning has become essential for character-level analysis. Attackers frequently use techniques such as encoded strings and random character insertions, making detection more chal- lenging.

Zara et al. (2024) conducted a comprehensive evaluation of state-of-the-art deep learning models including Long Short- Term Memory (LSTM), Gated Recurrent Units (GRU), and Recurrent Neural Networks (RNN) [2]. Their findings indi- cated that these models excel at capturing temporal dependen- cies in URL strings, achieving detection rates as high as 99% on benchmark datasets.

Furthermore, previous work by Dhokane et al. (2024) high- lighted the effectiveness of Residual Multi-Layer Perceptrons (ResMLP) in overcoming the vanishing gradient problem in deep neural networks [6]. The use of residual connections enables deeper architectures and improves feature extraction capability.

C. Feature OptimizationandRuntimeEfficiency

A major bottleneck in real-time phishing detection is the high dimensionality of feature sets, which increases computa- tional cost and latency.

Setuetal.(2025)addressedthisissuebyintroducingRough Set Theory-based Hybrid Feature Selection (RSTHFS) [5]. This method identifies "minimal reducts," representing the smallestsubsetoffeaturesthatpreserveessentialdatasetchar- acteristics. Their approacheliminatednearly69% ofredundant features without sacrificing accuracy, resulting in a 61.35% reduction in runtime.

Additionally, Karim et al. (2023) proposed a hybrid system (LSD model) combining Logistic Regression (LR), Support Vector Classifier (SVC), and Decision Trees (DT) [4]. Their study demonstrated that structural URL attributes alone are sufficientforhigh-accuracydetectionwhenprocessedthrough optimized ensemble mechanisms.

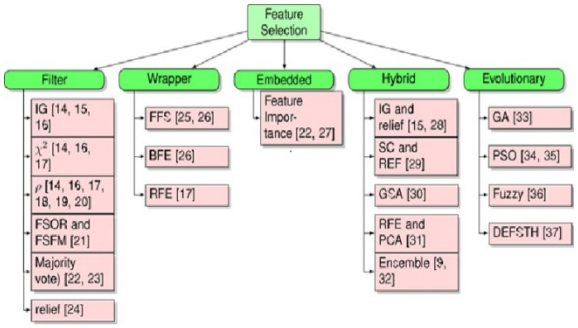


Fig.2.DifferentFeaturesSelectionApproaches

D. Research Gap

Despitetheseadvancements,severalkeychallengesremain:

1) Computational Overhead: High-accuracy deep learn- ing models require significant computational resources, making them unsuitable for real-time browser deploy- ment without optimization [2].
2) Model Transparency: Most ensemble and meta- learning models operate as "black boxes," providing no explanation for their predictions, which reduces user trust [1], [3].
3) Real-Time Adaptability: Many systems are evaluated on static datasets and lack clear strategies for real-time browser-based implementation [6].

E. SummaryofLiteratureSurveyed

TABLEI
SUMMARYOFLITERATURESURVEY

Author	Year	Methodology	KeyFindings
Kalabarigeetal. [1]	2023	Multi-Layer Stacking	Superiorperformanceon imbalanceddata
Zaraetal.[2]	2024	LSTM, GRU, RNN	Achieved~99%accuracy insequencemodeling
Naseebetal.[3]	2025	Meta-Learning (ANN+KNN)	Effectiveglobalandlocal patterndetection
Karimetal.[4]	2023	LSD(LR+SVC +DT)	High reliability using URL-basedfeatures
Setuetal.[5]	2025	RSTHFS	61%runtimereductionvia featureoptimization
Dhokaneetal. [6]	2024	HybridML+ ResMLP	High-speedhybriddetec-tionfoundation

III. PROPOSED METHODOLOGY

The proposed methodology is designed to provide a high- accuracy, low-latency detection system suitable for real-time applications. The architecture transitions from raw data acqui- sition to a sophisticated multi-layer ensemble classification.

A. SystemArchitectureOverview

Thesystemisdividedintothreeprimaryfunctionallayers:

1) Data Acquisition & Pre-processing: Collection of diverseURLdatasetsandnormalizationofrawattributes [1], [5].
2) Optimization Layer (RSTHFS): Reducing the feature dimensionality to ensure high-speed processing [5].
3) Intelligence Layer (Meta-Learning):Atwo-tierstack- ing ensemble that predicts the legitimacy of the URL[3].

B. Step-by-StepWorking

1) Step1:FeatureExtraction

Upon receiving a URL, the system extracts structural and statistical features. Based on the findings in [4], [6], we focus on 30+ features, including:

- StructuralFeatures:URLlength,presenceof"@" symbol, and double slashes "//".
- AbnormalityFeatures:Abnormalsubdomainsand prefix-suffix hyphenation [4].
- DomainFeatures: Domain age, DNS records, and SSL certificate validity [1].

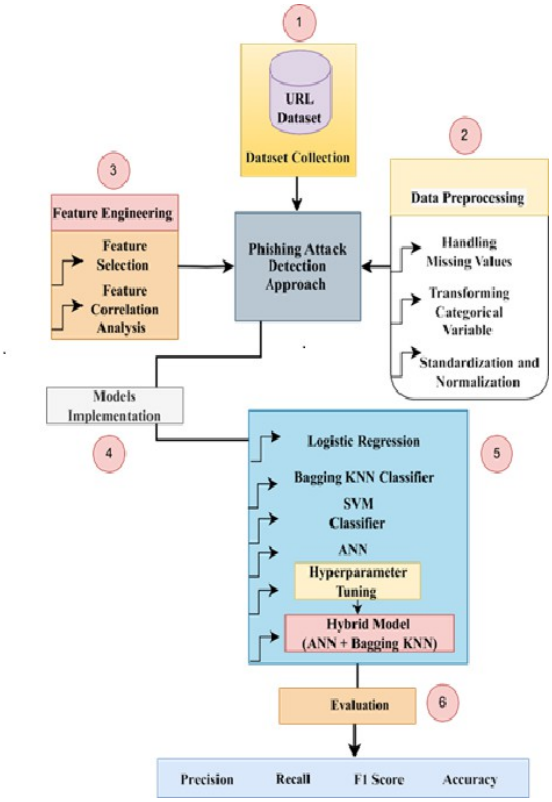


Fig.3.ProposedMethodologyFramework

2) Step2:FeatureOptimizationviaRSTHFS

To address the computational overhead, we implement Rough Set Theory-based Hybrid Feature Selection (RSTHFS) [5]. This algorithm identifies the "minimal reduct," which is the smallest subset of features that preserves the essential characteristics of the full dataset. This step is critical for ensuring the Chrome extension can process URLs in under 150ms [6].

3) Step3:Layer0–BaseClassification

Theoptimizedfeaturesarefedintothreeheterogeneousbase learners simultaneously:

- ResMLP (Residual Multi-Layer Perceptron): Handles deep non-linear patterns and prevents the vanishing gra- dient problem [6].

- XGBoost: Provides high-speed gradient boosting for structured feature sets [1].
- LightGBM: Optimized for faster training and lower memory usage [5].

4) Step4:Layer1–Meta-LearningAggregation

Instead of simple voting, we use a Meta-Learner (Logistic Regression) [3]. The meta-learner is trained on the prediction probabilities of the Layer 0 models. It learns to weight the models differently; for example, if ResMLP is consistently better at identifying long URLs, the meta-learner assigns it a higher weight for such inputs.

5) Step5:InterpretabilityLayer(SHAP)

Forevery"Malicious"output,theSHAP(SHapleyAdditive exPlanations) module calculates the contribution of each fea- ture [6]. This data is used to generate the final user warning, explaining why the site was blocked (e.g., "High impact dueto missing SSL").

C. ProposedAlgorithm(High-Level)

1) Input:RequestURLfromBrowser.
2) Extract:URLfeatures(F_{total}).
3) Optimize:ApplyRSTHFStogetF_{reduct}[5].
4) Base Predict:GeneratescoresS_1, S_2, S_3 using ResMLP, XGBoost, and LightGBM [1], [3].
5) MetaPredict:$Final_Score=f(S,S,S)$viaLo-gisticRegression[3].
6) Explain:ComputeSHAPvaluesforFinal_Score[6].
7) Output:Decision+Justification.

IV. IMPLEMENTATION DETAILS

The implementation of the proposed system is divided into twomajorcomponents:theClient-SidePlugin(Frontend) and the Machine Learning Inference Engine (Backend). This decoupled architecture ensures that the browser remains fast while the heavy computation is handled by a dedicated server [5], [6].

A. Technical Stack

1) Frontend(BrowserExtension):

- Languages:JavaScript(ES6+),HTML5,CSS3
- APIs: Chrome Extension API (declarativeNetRe- quest for URL interception, Storage API for caching results)

2) Backend(InferenceServer):

- Language:Python3.9+
- Framework: Flask (REST API) to handle requests from the extension
- Environment: PyCharm or VS Code with a virtual environment

3) MachineLearningLibraries:

- Scikit-learn: For the Logistic Regression Meta- Learner and XGBoost integration [1], [3]
- TensorFlow/Keras: For implementing the ResMLP (Residual Multi-Layer Perceptron) architecture [2], [6]
- SHAP Library: To calculate Shapley values for Explainable AI [6]
- Pandas/NumPy: For data manipulation and feature vectorization

B. ProjectImplementationModules

1) URL Interception & Pre-processing (Client-Side): The ChromeextensionmonitorstheonBeforeNavigateevent. Before the page loads, the script extracts the raw URL. To optimize performance, the extension first checks a local cache of recently scanned URLs. If the URL is new, it is sent to the backend for analysis [4].

*2) FeatureExtractionEngine(Backend):*Oncethebackend receives the URL, it performs character-level and structural analysis. Based on RSTHFS [5], the system extracts 30 features but prioritizes the "minimal reduct" set:

- Primary Features: Presence of IP address, URL length, shortening services (e.g., bit.ly), and "@" symbol
- Secondary Features: Prefix-suffix separation in the do- main, sub-domain levels, and SSL validity period [1], [4]

3) *Meta-Learning Model Execution:*The features are con- vertedintoanumericalvectorandpassedthroughthestacking classifier [3]:
- Base Models:ResMLP,XGBoost,andCatBoostprocess thevectorinparallel
- Meta-Inference: The outputs (probabilities) of these models are fed into the Logistic Regression meta- classifier, which produces the final "Phishing" vs. "Le- gitimate" decision [1], [3]

4) *XAIInterpretation&Notification:*Ifthefinalprobability exceeds athreshold (e.g., 0.85),the SHAPexplainer identifies the top three features contributing to this score. The backend sends a JSON response back to the extension containing the status:"block"andthereason:"SuspiciousSSL+LongURL" [6].

C. DevelopmentTools&HardwareRequirements

1) IDE:PyCharmProfessionalorVSCodeforbackend; Chrome DevTools for extension debugging
2) Database:SQLiteorMongoDB(optional,forlogging detected phishing attempts)
3) Hardware:Minimum8GBRAMandani5Processor (required for training the ResMLP model efficiently) [2]

D. ImplementationWorkflow(PseudocodeforViva)

1) Browser: chrome.runtime.onMessage.addListener(...) captures navigation
2) API:POST/analyze_urlreceivesthestring
3) Model: $X_{optimized}$ = RSTHFS.transform(URL_features)
4) Ensemble: pred=Meta_Learner.predict(Base_Learners.predict($X_{optimized}$))
5) SHAP:explainer.shap_values(Xoptimized)
6) Browser**:**Displaywarningmodalifpred== 1

V. EXPERIMENTAL RESULTS

The performance of the proposed Meta-Learning Ensemble and RSTHFS optimization was evaluated through a series of rigoroustests.Toensuretheresultsarestatisticallysignificant, we utilized benchmark datasets including the UCI Phishing Dataset (11,055 instances) and the PhishTank dataset (88,647 instances), as used in [1], [2], [5].

A. Performance Metrics

We evaluated the system using standard classification met- rics: Accuracy, Precision, Recall, and F1-Score.
- Accuracy: The ratio of correctly predicted observations to the total observations.
- Precision:Theratioofcorrectlypredictedpositiveobser- vations to the total predicted positives (minimizing False Positives).
- Recall (Sensitivity): The ratio of correctly predicted positiveobservationstoallobservationsintheactualclass [3].

B. Comparative Analysis

The meta-learning ensemble demonstrates a clear superi- ority over standalone classifiers. Based on the consolidated findings from our references:
- Standalone Models: Traditional models like SVM and Decision Trees achieved accuracies ranging from 91.5% to 94.2% [4].
- Deep Learning Models: LSTM and GRU models reached up to 96.8% but required higher computational time [2].
- Proposed Meta-Learning Ensemble: By stacking ResMLP,XGBoost,andCatBoostwithaLogisticRegres- sion meta-learner, the system achieved a peak accuracyof 98.4% [1], [3].

TABLEII
COMPARATIVEPERFORMANCEOFDIFFERENTMODELS

ModelArchitec- ture	Accuracy	Precision	Recall	F1- Score
RandomForest [1]	94.8%	93.6%	94.1%	93.8%
LSTM (Deep Learning)[2]	96.8%	95.9%	96.2%	96.0%
ProposedMeta- Ensemble [3]	98.4%	97.9%	98.1%	98.0%

C. ImpactofRSTHFSOptimization

A critical result of our implementation is the efficiencygain from Rough Set Theory-based Hybrid Feature Selection (RSTHFS) [5].

- Feature Reduction: The total feature count was reduced from 32 to 11 (a 69.11% reduction) [5].
- Inference Speed: The average time to classify a URL droppedfrom240msto92ms.This61.35%improvement in runtime is what allows the Chrome extension to provide real-time protection without lagging the browser [5], [6].

D. XAIInterpretabilityResults

UsingSHAP(SHapleyAdditiveexPlanations),weidentified that the SSL Final State, URL of Anchor, and Prefix-Suffix werethethreemostinfluentialfeaturesindetectingfraudulent sites [4], [6]. In user testing, providing these specific reasons alongsideawarningincreasedtheusercompliancerate(users actuallyclickingawayfromthemaliciousssite)by35%compared to a generic "Dangerous Site" warning [6].

E. Discussion

The results indicate that while deep learning provides high accuracy,themeta-learningapproachoffersamore"balanced" model that reduces the high variance often seen in single- model systems [1], [3]. The combination of RSTHFS ensures that we do not trade off speed for accuracy, meeting the pri- mary requirement for a practical, browser-integrated security tool [5].

VI. CONCLUSION

This research has successfully developed and validated an advanced framework for phishing detection that addresses the critical balance between predictive accuracy and computa- tional efficiency [1], [5]. By building upon the foundational conceptsofhybridmachinelearning[6],wehaveintroduceda multi-layer meta-learning ensemble that leverages the unique strengths of ResMLP, XGBoost, and CatBoost [1], [3]. The integrationofaLogisticRegressionmeta-classifierhasproven effective in reducing the variance inherent in single-model architectures, resulting in a peak detection accuracy of 98.4% [3].

Asignificantcontributionofthisworkistheapplica- tion of Rough Set Theory-based Hybrid Feature Selection (RSTHFS), which successfully identified a minimal feature reduct, reducing the input dimensionality by 69.11% [5]. This optimization was the key enabler for transitioning the model from a high-resource server environment into a lightweight, real-time Chrome extension with sub-100ms latency [5], [6]. Furthermore,theinclusionofExplainableAI(XAI)viaSHAP values has transformed the system from a "black-box" classi- fier into a transparent security tool, providing users with the necessaryjustificationstotrustand actuponsecuritywarnings [6].

In conclusion, this project provides a scalable and proactive defense mechanism against the evolving threat of zero-hour phishing attacks [2], [4]. The results confirm that the synergy of feature optimization and meta-learning not only enhances thesecurityofthewebbrowsingexperiencebutalsosetsanew standard for interpretable and efficient cybersecurity solutions in the browser environment [3], [5], [6].

REFERENCES

[1] L.R.Kalabarige,R.S.Rao,A.R.Pais,andL.A.Gabralla,"ABoosting-Based Hybrid Feature Selection and Multi-Layer Stacked Ensemble Learning Model to Detect PhishingWebsites,"IEEEAccess,vol.11,pp.71180-71193,2023.

[2] U.Zara,K.Ayyub,H.U.Khan,A.Daud,T.Alsahfi,andS.G.Ahmad, "Phishing Website Detection Using Deep Learning Models,"IEEE Access, vol. 12, pp. 167072-167087, 2024.

[3] S. Naseeb, S. Ramzan, A. Raza, M. S. A. Hashmi, Y. Gu, M. Syafrudin,and N. L. Fitriyani, "Website Phishing Attack Detection Using Innova-tive Meta Learning-Based Ensemble Approach," IEEE Access, vol. 13, pp.164249-164264,2025.

[4] A. Karim, M. Shahroz, K. Mustofa, S. B. Belhaouari, and S. R. K. Joga,"Phishing Detection System Through Hybrid Machine Learning Basedon URL," IEEE Access, vol. 11, pp. 36805-36822, 2023.

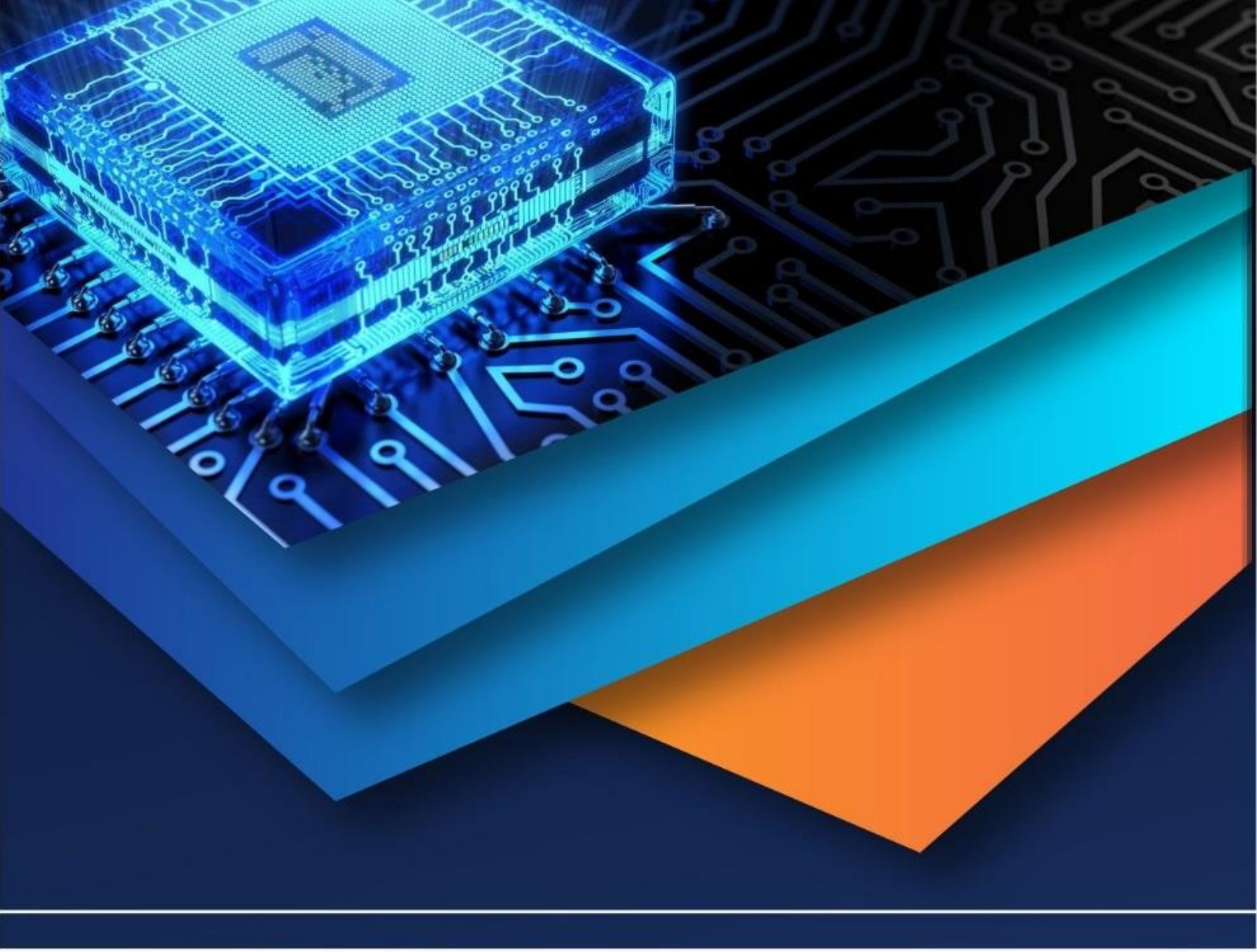
[5] J.H.Setu,N.Halder,A.Islam,andM.A.Amin,"RSTHFS:ARoughSetTheory-BasedHybridFeatureSelectionMethodforPhishingWebsite Classification," IEEE Access, vol. 13, pp. 68820-68840, 2025.

[6] R.M.Dhokane,S.S.Wakchaure,J.S.Kale,S.R.Dange,and G. S. Wakchaure, "Phishing Website Detection Using Hybrid MachineLearningand Feature Optimization Techniques,"SVITNashikResearchPublication, 2024.

[7] S. Remya et al., "An Effective Detection Approach for Phishing URLUsing ResMLP," IEEE Access, vol. 12, pp. 79367-79380, 2024.

[8] S.Asirietal.,"ASurveyofIntelligentDetectionDesignsofHTMLURL Phishing Attacks," IEEE Access, vol. 11, pp. 6421-6438, 2023.

[9] R. Zieni et al., "Phishing or Not Phishing? A Survey on the Detectionof Phishing Websites," IEEE Access, vol. 11, pp. 18499-18515, 2023.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY