CLASSIFICATION 0F PHISHING URL ATTACKS USING RANDOM FOREST ALGORITHM BASED ON FEATURE IMPORTANCE

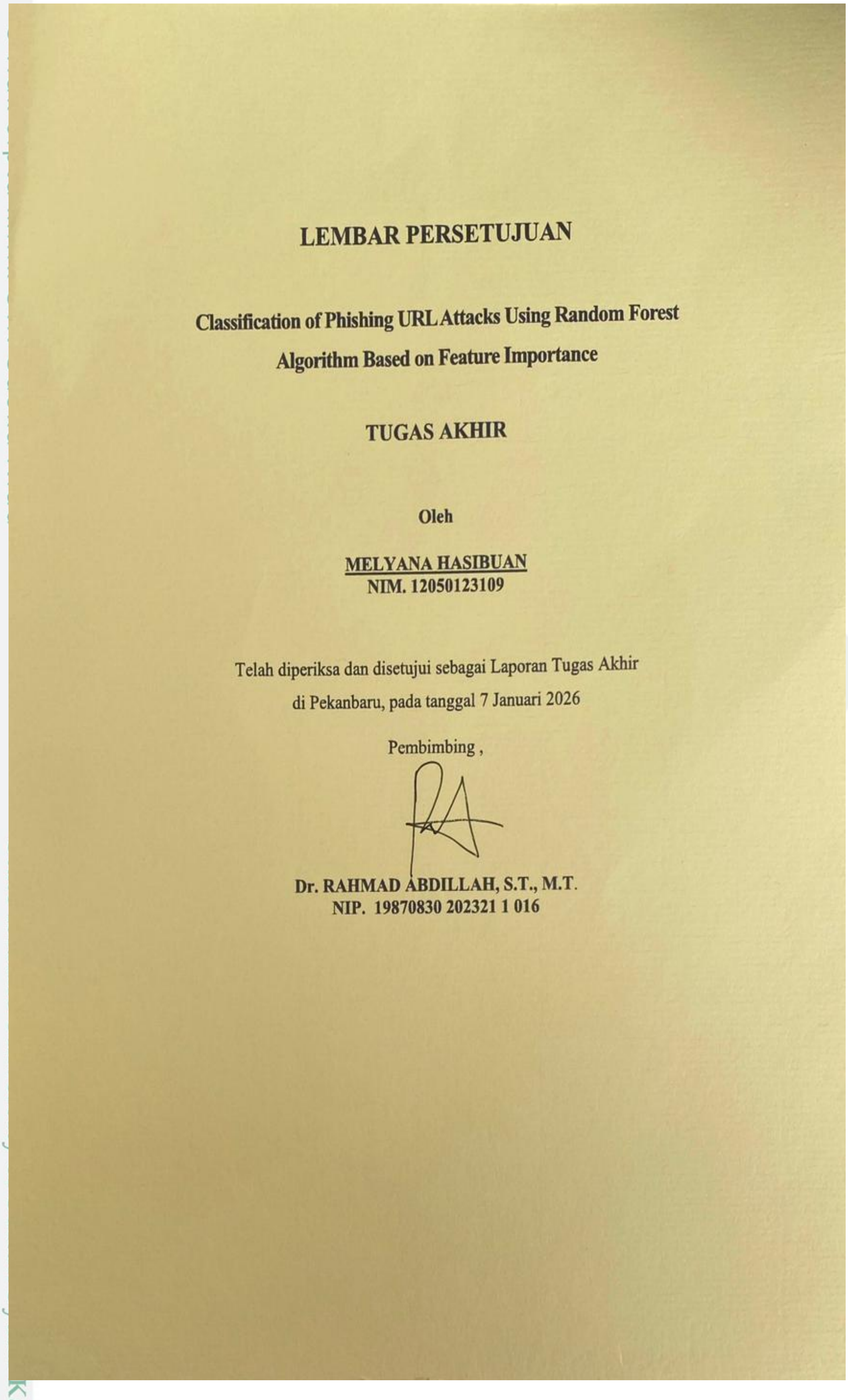
TUGAS AKHIR

Disusun Sebagai Salah Satu Syarat
Untuk Memperoleh Gelar Sarjana Teknik
Pada Jurusan Teknik Informatika

Oleh
MELYANA HASIBUAN
NIM. 12050123109

UIN SUSKA RIAU

FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU
PEKANBARU
2026



LEMBAR PERSETUJUAN

Classification of Phishing URL Attacks Using Random Forest Algorithm Based on Feature Importance

TUGAS AKHIR

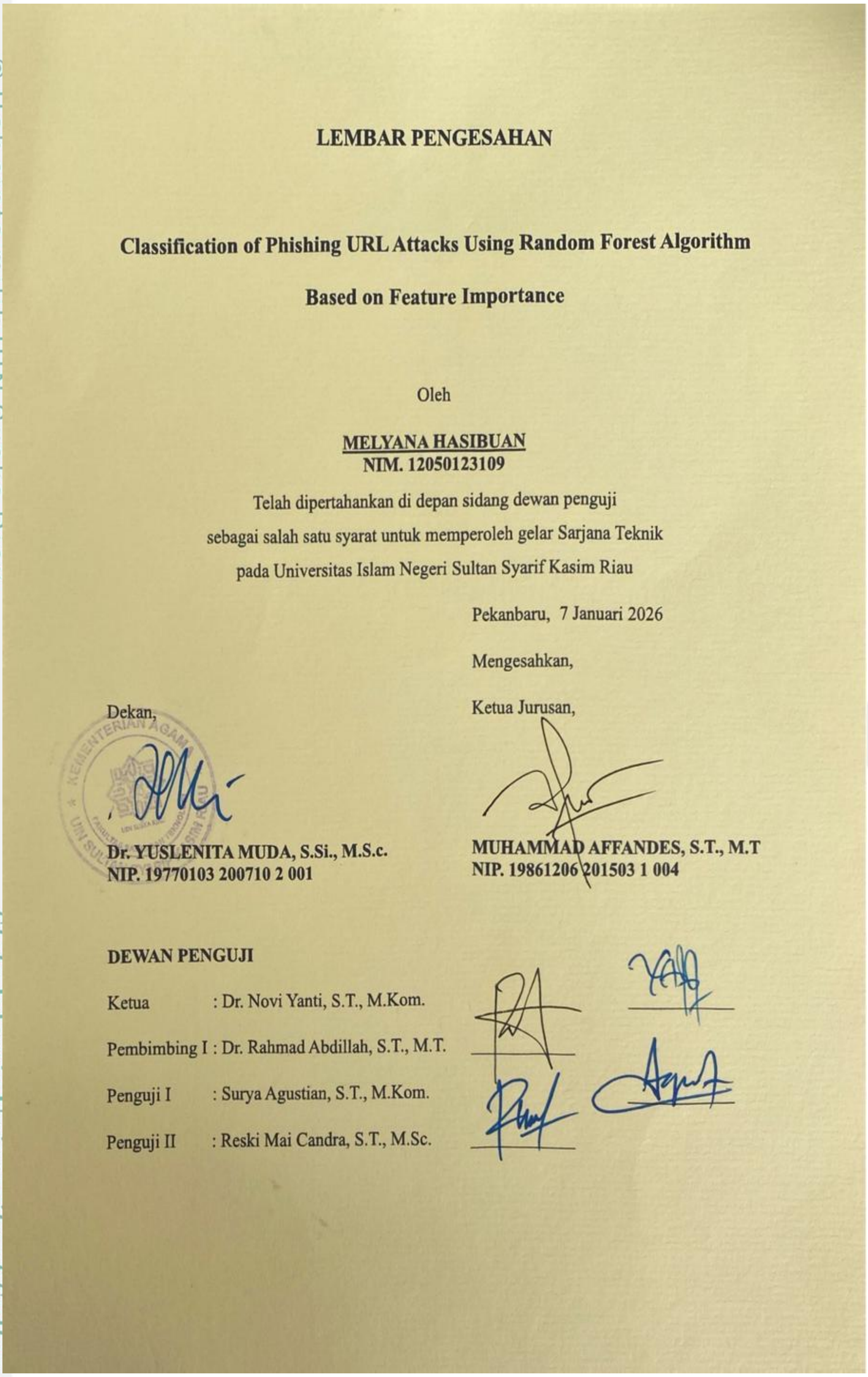
Oleh

MELYANA HASIBUAN
NIM. 12050123109

Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir
di Pekanbaru, pada tanggal 7 Januari 2026

Pembimbing ,

Dr. RAHMAD ABDILLAH, S.T., M.T.
NIP. 19870830 202321 1 016



LEMBAR PENGESAHAN

Classification of Phishing URL Attacks Using Random Forest Algorithm Based on Feature Importance

Oleh

MELYANA HASIBUAN
NIM. 12050123109

Telah dipertahankan di depan sidang dewan penguji
sebagai salah satu syarat untuk memperoleh gelar Sarjana Teknik
pada Universitas Islam Negeri Sultan Syarif Kasim Riau

Pekanbaru, 7 Januari 2026

Mengesahkan,

Dekan,

KEMENTERIAN AGAMA

Dr. YUSLENITA MUDA, S.Si., M.S.c.
NIP. 19770103 200710 2 001

Ketua Jurusan,

MUHAMMAD AFFANDES, S.T., M.T
NIP. 19861206 201503 1 004

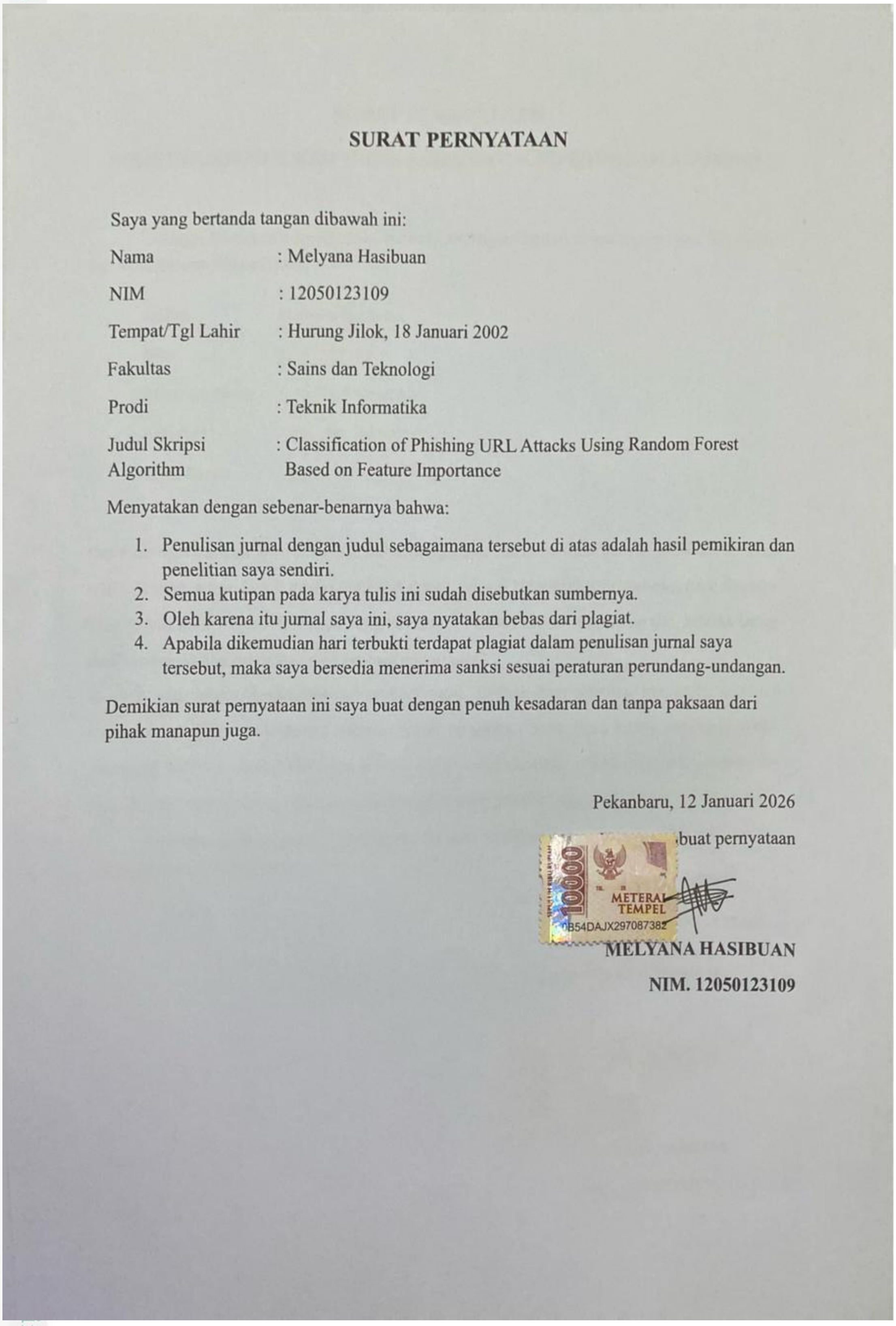
DEWAN PENGUJI

Ketua : Dr. Novi Yanti, S.T., M.Kom.

Pembimbing I : Dr. Rahmad Abdillah, S.T., M.T.

Penguji I : Surya Agustian, S.T., M.Kom.

Penguji II : Reski Mai Candra, S.T., M.Sc.



SURAT PERNYATAAN

Saya yang bertanda tangan dibawah ini:

Nama : Melyana Hasibuan

NIM : 12050123109

Tempat/Tgl Lahir : Hurung Jilok, 18 Januari 2002

Fakultas : Sains dan Teknologi

Prodi : Teknik Informatika

Judul Skripsi : Classification of Phishing URL Attacks Using Random Forest Algorithm Based on Feature Importance

Menyatakan dengan sebenar-benarnya bahwa:

1. Penulisan jurnal dengan judul sebagaimana tersebut di atas adalah hasil pemikiran dan penelitian saya sendiri.
2. Semua kutipan pada karya tulis ini sudah disebutkan sumbernya.
3. Oleh karena itu jurnal saya ini, saya nyatakan bebas dari plagiat.
4. Apabila dikemudian hari terbukti terdapat plagiat dalam penulisan jurnal saya tersebut, maka saya bersedia menerima sanksi sesuai peraturan perundang-undangan.

Demikian surat pernyataan ini saya buat dengan penuh kesadaran dan tanpa paksaan dari pihak manapun juga.

Pekanbaru, 12 Januari 2026

buat pernyataan

10000 METERAI TEMPEL 0B54DAJX297087382

MELYANA HASIBUAN

NIM. 12050123109

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa:

1. Tugas Akhir ini dengan judul "Classification of Phishing URL Attacks Using Random Forest Algorithm Based on Feature Importance" adalah gagasan asli dari saya sendiri dan belum pernah dijadikan Tugas Akhir atau sejenisnya di Universitas Islam Negeri Sultan Syarif Kasim Riau maupun di perguruan tinggi lain.
2. Dalam Tugas Akhir ini TIDAK terdapat karya atau pendapat yang telah dipublikasikan orang lain, kecuali tertulis dengan jelas dan dicantumkan sebagai referensi di dalam Daftar Pustaka.
3. Dalam Tugas Akhir ini TIDAK terdapat penggunaan Kecerdasan Buatan Generatif (Generative AI) yang bertentangan dengan ketentuan dan peraturan yang berlaku.
4. Saya bersedia menerima sanksi sesuai ketentuan yang berlaku apabila di kemudian hari terbukti bahwa Tugas Akhir ini melanggar kode etik maupun peraturan yang berlaku, termasuk plagiat ataupun pelanggaran hak cipta.

Demikianlah pernyataan ini dibuat dengan sebenarnya.

Pekanbaru, 12 Januari 2026
Yang membuat pernyataan,

MELYANA HASIBUAN
NIM.12050123109

LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

Tugas Akhir yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum dengan ketentuan bahwa hak cipta pada penulis. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan seizin penulis dan harus disertai dengan kebiasaan ilmiah untuk menyebutkan sumbernya.

Penggandaan atau penerbitan sebagian atau seluruh Tugas Akhir ini harus memperoleh izin dari Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan yang meminjamkan Tugas Akhir ini untuk anggotanya diharapkan untuk mengisi nama, tanda peminjaman dan tanggal pinjam.

LEMBAR PERSEMBAHAN

بِسْمِ اللَّهِ الرَّحْمَٰنِ الرَّحِيمِ

Alhamdulillahi robbil’alamin..

Tugas akhir ini penulis persembahkan sebagai bentuk semangat, usaha, serta ungkapan cinta dan kasih sayang kepada orang-orang terpenting dalam hidup penulis. Dengan ketulusan hati dan rasa terima kasih yang mendalam, tugas akhir ini penulis persembahkan kepada:

1. Kedua orang tua tercinta, Bapak Mara Bangun Hasibuan dan Ibu Yusra Harahap, serta adik-adik ku tersayang, juga seluruh keluarga besar penulis yang telah memberikan dukungan moral, materil, serta doa dan restu, sehingga penulis dapat menempuh pendidikan hingga jenjang S1 di Jurusan Teknik Informatika, UIN Sultan Syarif Kasim Riau.
2. Dosen pembimbing, Bapak Dr. Rahmad Abdillah, S.T., M.T., yang telah memberikan bimbingan, arahan, dan motivasi hingga tugas akhir ini dapat terselesaikan dengan baik
3. Seluruh dosen pengajar yang telah membimbing dan mendidik penulis dengan penuh kesabaran dan keikhlasan, sehingga ilmu yang diperoleh selama masa perkuliahan dapat menjadi bekal yang bermanfaat di masa depan.
4. Teman-teman seperjuangan di Program Studi Teknik Informatika, UIN Sultan Syarif Kasim Riau, atas kebersamaan dan dukungan selama menempuh perjalanan akademik.

Semoga tugas akhir ini dapat memberikan manfaat bagi para pembaca. Aamiin ya Rabbal ‘Alamiin.

Bit-Tech (Binary Digital - Technology)

Vol.8, No.2, December 2025
Available online at: http://jurnal.kdi.or.id/index.php/bt

Classification of Phishing URL Attacks Using Random Forest Algorithm Based on Feature Importance

Melyana Hasibuan[1)*], Rahmad Abdillah[2)], Surya Agustian[3)], Reski Mai Candra[4)]

[1)2)3)4)] *Informatics Engineering, Faculty of Science and Technology, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia*

[1)]12050123109@students.uin-suska.ac.id
[2)]rahmad.abdillah@uin-suska.ac.id
[3)]surya.agustian@uin-suska.ac.id
[4))]reski.candra@uin-suska.ac.id

Article history:

Received 02 Dec 2025;
Revised 05 Dec 2025;
Accepted 09 Dec 2025;
Available online 10 Dec 2025

Keywords:

Feature Importance
Machine Learning
Phishing URL Detection
Random Forest
URL Security

Abstract

The development of information technology and increasing digital activities have made URL-based phishing threats more complex and difficult to detect. Phishing attacks target not only individuals but also organizations, requiring detection systems that are accurate, efficient, and capable of handling high-dimensional data. Machine learning approaches, particularly Random Forest, have been widely applied for phishing detection; however, further evaluation is needed regarding the role of feature selection in improving efficiency without reducing performance. This study aims to evaluate the performance of the Random Forest algorithm for phishing URL detection and to analyze the impact of feature selection based on feature importance. This research adopts the Knowledge Discovery in Databases (KDD) framework, including data selection, preprocessing, feature selection, modeling, and evaluation stages. The PhiUSIIL-2024 dataset is used, with two modeling scenarios: Random Forest using all features (RF Full) and Random Forest using the top 30 features selected through feature importance (RF Top-30). Model performance is evaluated using accuracy, precision, recall, and F1-score metrics under different data split ratios. The experimental results show that both models achieve very high and stable classification performance, with evaluation metrics close to or reaching 100%. The RF Top-30 model maintains performance comparable to the RF Full model despite using fewer features. This study concludes that feature importance-based feature selection effectively simplifies the Random Forest model without sacrificing performance, making it suitable for efficient URL phishing detection systems.

I. INTRODUCTION

The Internet has become the primary infrastructure that supports a wide range of digital activities, including communications, business, education, and financial and government services [1]. Along with the rapid adoption of cloud computing, mobile applications, digital payment systems, and web-based platforms, internet dependency has increased significantly across both individual and organizational sectors. The massive increase in internet usage is directly followed by an increase in cybersecurity threats with increasingly complex and unpredictable attack patterns [2]. This condition makes information security a critical aspect that must be considered in the development and utilization of modern digital systems. Various types of cyber attacks often occur, including malware, ransomware, denial of service (DoS), and social engineering techniques that exploit user weaknesses. Of these threats, phishing is one of the most dangerous attacks because it utilizes psychological manipulation through fake sites that resemble official websites to steal sensitive user data, such as account credentials and financial information [1], [3]. In recent years, phishing attacks have evolved by leveraging shortened URLs, dynamically generated domains, HTTPS spoofing, and domain obfuscation techniques, making them increasingly

[*] Melyana Hasibuan

ISSN 2622-2728 (online) 2622-271X (print) © 2018 The Authors. Published by Komunitas Dosen Indonesia.

doi: 10.32877/bt.v8i2.3511

difficult to detect using traditional approaches. The impact of phishing attacks not only causes financial losses, but also lowers the level of public trust in digital services [4].

A variety of methods have been developed to detect cyber threats, including phishing, malware, ransomware, and social engineering [5]. Conventional approaches such as blacklist and whitelist have significant limitations because they are only able to detect previously known attacks, making them less effective in dealing with new attacks or zero-day attacks [6]. These limitations drive the need for detection systems that are more adaptive, proactive, and capable of recognizing dynamic attack patterns. Therefore, the use of machine learning (ML) is a promising solution due to its ability to learn complex patterns from data and perform automatic classification [7]–[9]. As phishing techniques continue to adapt to modern web technologies, machine learning-based detection systems must also be continuously evaluated using up-to-date datasets to ensure their effectiveness in real-world scenarios.

Prior research has demonstrated the efficiency of machine learning methods in detecting phishing, especially in URL classification. Algorithms such as Decision Tree, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Random Forest (RF) have been widely used and compared in performance [10], [11]. Among these algorithms, Random Forest consistently showed superior and stable performance in various phishing detection studies. This is because it can handle high-dimensional data, lower the chance of overfitting, and aggregate the output of several decision trees to provide predictions that are more accurate. Random Forest is an ensemble learning method that works by building hundreds of decision trees and combining the results to obtain more accurate and stable predictions [3].

Many previous studies have specifically highlighted the advantages of excellent RF. For example, a study by [12] managed to achieve an accuracy of 98% with high precision, recall, and F1-score values. Another study by [13] reported an accuracy of 95% on a parameter-adjusted Random Forest model. Meanwhile, comparative research by [11] showed that Random Forest outperformed Naïve Bayes, Decision Tree, and SVM with an accuracy of 90.77%. These findings reinforce Random Forest's position as one of the most reliable algorithms in the context of URL-based phishing detection.

However, most previous studies still have important limitations that need to be noted. One of the main drawbacks is the use of relatively outdated datasets, so they no longer represent the current patterns and characteristics of phishing attacks [4], [14]. Phishing tactics are always changing., models trained using old data risk losing their relevance when applied to real-world conditions. In addition, some studies focus more on achieving high accuracy values without evaluating the stability of the model against parameter variations or the efficiency of the number of features used. In many cases, feature selection is treated as a secondary process or not analyzed explicitly, even though the number and relevance of features play a critical role in model efficiency, interpretability, and generalization performance. In fact, in a cybersecurity detection system, the efficiency and generalization of the model are as important as the value of accuracy alone.

Based on these limitations, a clear research gap can be identified: there is a lack of studies that explicitly analyze the role of feature selection particularly feature importance-based selection in optimizing Random Forest performance for phishing URL detection using a truly state-of-the-art dataset. Some recent studies have indeed started using modern datasets, but most have shifted to deep learning approaches that have high computational complexity and are less interpretive [15]. This trend often overlooks the potential of classical ensemble methods, such as Random Forest, to achieve comparable performance with significantly lower computational cost and higher model interpretability when combined with appropriate feature selection techniques. This opens up an opportunity to re-evaluate the potential of Random Forest as a simpler, more efficient, and still competitive method.

Based on this background, this study focuses on the application of the Random Forest algorithm for the classification of phishing URL attacks using the modern PhiUSIIL-2024 dataset. This study not only evaluates the performance of Random Forest with all features, but also analyzes the effectiveness of feature selection based on feature importance using the top 30 features. In addition, this study examines the stability of model performance against variations in the training-test data ratio as well as the main parameters of Random Forest, namely the number of trees (n_estimators) and the maximum depth of trees (max_depth).

The primary goal of this research is to develop and test a Random Forest-based phishing URL categorization model that is high-performing, stable, and efficient. This study specifically aims to: (1) assess Random Forest's performance on the most recent URL phishing dataset; (2) examine how feature selection based on feature importance affects model performance; and (3) compare the stability of classification results in different parameter scenarios and data sharing ratios. It is believed that the findings of this study will contribute to science by deepening our understanding of Random Forest's efficacy in contemporary phishing detection and serving as the foundation for the creation of a dependable and successful automated phishing detection system.

II. Literature Review

Numerous studies have examined the use of machine learning for phishing detection, with findings that consistently highlight the effectiveness of certain algorithms, while also identifying ongoing challenges. In general, it is concluded that Random Forest (RF) is one of the most reliable algorithms for this task. The main similarity in the literature is the recognition of the superiority of RF performance over other classification methods. Studies by [8], [10] reported that RF achieved the highest accuracy, 98.35% and 99.45%, respectively, surpassing algorithms such as KNN and SVM. These findings are supported by various other studies that also position RF as a superior model in the context of URL and phishing email detection [14], [16], [17], which collectively confirm the consistency of RF's performance across different phishing data scenarios and types.

Despite its excellent performance, most previous studies have focused on achieving accuracy values alone, without conducting an in-depth analysis of model efficiency and performance stability to parameter variations. In addition, many studies use datasets that do not fully reflect current phishing attack patterns, raising concerns about the model's generalization capabilities when implemented in dynamic real-world environments. This limitation is an important issue, considering that the characteristics of phishing continue to evolve as technology and user behavior changes.

Recent research shows that the RF algorithm remains one of the most effective methods in handling high-dimensional data in phishing detection cases, even when compared to deep learning-based approaches. In studies using the modern PhiUSIIL dataset, the Fully Connected Neural Network (FCNN) model was reported to be able to achieve an accuracy of 99.38%. However, in the same study, Random Forest also showed competitive performance with an accuracy rate of 98.69% despite having a much lower model complexity [15]. These findings indicate that Random Forest still has strong competitiveness, especially in terms of computational efficiency and ease of model interpretation.

However, these studies have not explicitly evaluated the optimization potential of Random Forest through a feature-importance-based feature selection approach on a truly state-of-the-art URL phishing dataset. Some studies focus more on comparisons between algorithms or the application of complex models, without exploring the extent to which feature reductions can maintain or even improve model performance. More specifically, prior research has not systematically investigated how integrating feature importance-based feature selection within the Random Forest framework affects classification performance, model stability, and computational efficiency in phishing URL detection tasks. In fact, the right selection of features has the potential to reduce model complexity, speed up computational time, and increase interpretability without sacrificing accuracy. Based on this explicit research gap, there is still limited empirical evidence that demonstrates whether Random Forest models optimized through feature importance-based feature selection can achieve performance comparable to or better than full-feature models when applied to modern, large-scale phishing datasets. This gap becomes particularly critical in the context of URL phishing detection, where high-dimensional feature spaces can negatively impact computational efficiency and model interpretability if not properly managed.

Departing from these findings, this study is positioned to fill the research gap by evaluating the performance of the Random Forest algorithm which is optimized through feature selection based on feature importance in the modern PhiUSIIL-2024 dataset. In contrast to previous studies, this study not only assessed the accuracy of the model, but also analyzed the stability of Random Forest's performance on various parameter scenarios and data sharing ratios. The evaluation was conducted comprehensively using accuracy, precision, recall, and F1-score metrics to ensure that the resulting model was not only numerically superior, but also reliable in detecting URL phishing threats under current conditions.

III. Methods

The research methodology is a systematic framework that outlines the steps to be carried out in the research. This study adopts the Knowledge Discovery in Databases (KDD) framework as a systematic approach to ensure that each stage of research, from data collection and selection, pre-processing, transformation, to evaluation of classification results, is done in a methodical and repeated way. The KDD framework has been widely used in data mining-based research because it provides a well-defined sequence of steps to extract patterns and knowledge from big data through the processes of selection, cleaning, transformation, data mining, and interpretation and evaluation of results [18]. The flow of the research methodology to be carried performed is shown in Figure 1.

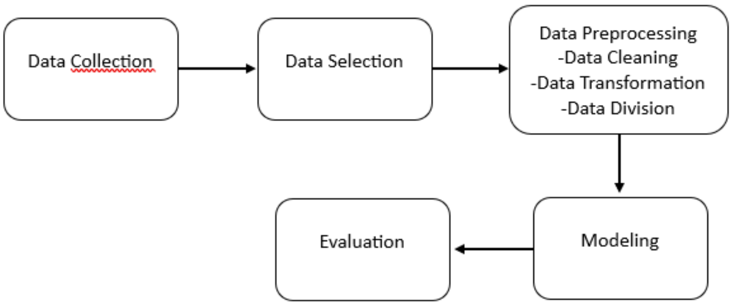


Fig. 1 Research Methodology

A. Data Collection

This stage focuses on collecting secondary data that will be the main material of the research. The data source used comes from a public dataset titled Phishing URL Websites Dataset (PhiUSIIL) developed by Prasad and Chandra [20]. which was acquired from the Kaggle platform and updated in 2024. This dataset can be accessed via the link: https://www.kaggle.com/datasets/kaggleprollc/phishing-url-websites-dataset-phiusiil?resource=download. This dataset is large-scale with a total of 235,795 URL data. Each piece of data is represented by 56 columns, consisting of 1 identification column (text URL), 54 feature columns that represent the various characteristics of that URL, and 1 target column (Label). This data is divided into two classes, namely 134,850 legitimate URLs and 100,945 phishing URLs, which will be classified into Phishing (Label 1) and Non-Phishing (Label 0) target labels. The initial dataset can be seen in Table 1 below:

TABLE 1
INITIAL DATASET

No	Filename	URL	URL Length	Domain	Domain Length	TLD	…	Label
1	521848.txt	https://www.southbankmosaics.com	31	www.southbankmosaics.com	24	com	…	1
2	31372.txt	https://www.uni-mainz.de	23	www.uni-mais.de	16	ce	…	1
3	mw42508.txt	http://www.teramil.com	22	www.teramill.com	16	com	…	0
4	151578.txt	https://www.rewildingargentina.org	33	www.rewildingargentina.org	26	org	…	1
5	mw16985.txt	http://www.fO519141.xsph.ru	26	www.fO519145.sph.ru	20	ru	…	0

B. Data Selection

At this stage, the selection of variables to be used in the modeling process is performed to determine the most relevant attributes for classification. The URL, Filename, and Domain fields containing raw textual data are removed because they function solely as unique identifiers for each data instance. These attributes do not provide direct predictive value that can be effectively processed by the classification algorithm. Consequently, the modeling process is conducted using 52 feature variables and one target class.

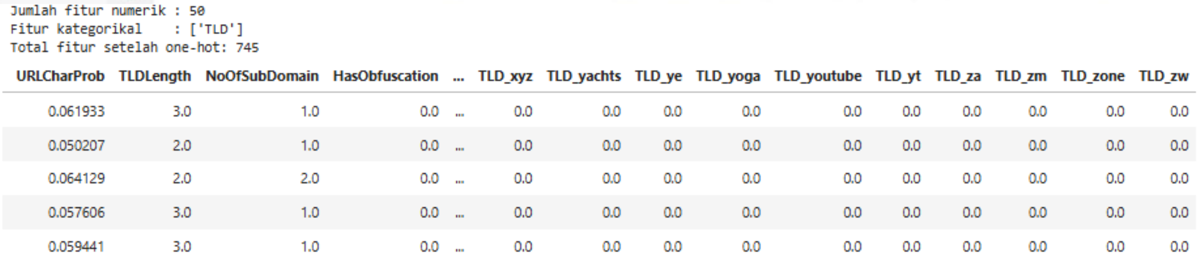
C. Pre-Processing Data

This stage aims to clean the data of potential issues that could interfere with the model's performance. The main step that will be taken is data cleaning. The dataset will be thoroughly examined to identify Duplicate data with missing values (NaN). If duplicate data is found, the line will be removed to maintain data integrity. If there is a missing value, the imputation method will be applied by replacing it using the mode value (the most frequently occurring value) in that column, especially for categorical features.

All information will be transformed into a completely numerical representation so that Random Forest's algorithms can process it, with a primary focus on encoding categorical features. Most of the 52 attributes are already in numerical or binary format (0/1). However, TLD variables are still in text (categorical) format. Therefore, for these categorical variables, the One-Hot Encoding technique will be applied to convert each unique category into new binary columns. As Table 2 below illustrates:

TABLE 2
COMPARISON OF VARIABLES BEFORE AND AFTER ONE HOT ENCODING

Before One Hot Encoding	After One Hot Encoding
TLD	TLD_.com TLD_.net TLD_.org … TLD_.zw



```
Jumlah fitur numerik : 50
Fitur kategorikal    : ['TLD']
Total fitur setelah one-hot: 745
```

URLCharProb	TLDLength	NoOfSubDomain	HasObfuscation	...	TLD_xyz	TLD_yachts	TLD_ye	TLD_yoga	TLD_youtube	TLD_yt	TLD_za	TLD_zm	TLD_zone	TLD_zw
0.061933	3.0	1.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.050207	2.0	1.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.064129	2.0	2.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.057606	3.0	1.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.059441	3.0	1.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Fig. 2 Attribute confusion after TLD encoding

Based on the results of the encoding process shown in Figure 2, the number of attributes in the dataset has increased significantly from 51 to 745 attributes. This improvement is mainly due to the Top-Level Domain (TLD) attribute having 695 different instances, resulting in a huge number of features after the implementation of the one-hot encoding technique. It should be noted that at this stage there is no data normalization process because the Random Forest algorithm is a decision tree-based method that is insensitive to the difference in value scales between features. In addition, most of the features in the dataset are binary (0/1) after the encoding process, so normalization does not provide additional benefits and instead adds complexity. Therefore, after the encoding process, the dataset is considered ready to proceed to the next stage of modeling.

Two datasets will be created from the cleaned and modified dataset: training data and test data. This division will be done for two different ratio scenarios: 90:10 and 80:20 [19]. For each scenario, the division will use the stratified splitting technique. The reason for using this technique is to ensure that the proportion of phishing and non-phishing classes remains balanced across both datasets, according to the distribution in the original dataset. This is crucial to prevent imbalances in the evaluation and ensure the model is tested on data that represents real-world problems.

D. Modeling

The KDD process's central phase is this one. Each partitioning scenario's training data will be subjected to the Random Forest method. Numerous decision trees (n_estimators) will be individually constructed by this procedure. Each tree will only take into account a random subset of characteristics on each of its node separations after it's trained on a random sample of the training data (bagging approach). The goal of this ensemble technique is to create robust, accurate, and highly resistant overfitting models.

Although Random Forest can naturally handle many features, the use of 52 features directly risks including irrelevant features (noise) that can affect the computational time and accuracy of the model. Therefore, the feature selection stage will be carried out to identify the subset of features that are most influential and have the potential to improve model performance. The method used is Feature Importance which is extracted directly from the Random Forest model itself. This study will reduce the features to 30 (Top-30) based on the previous research i.e. research by [22]. The advantage of this method is that the selection of features is based on how much each feature contributes to the voting process and the reduction of impurity (Gini) during the model being trained. Here is the Equation (1) for calculating Gini.

$$Gini = 1 - \sum_{i=1}^{C} (P_i)^2 \tag{1}$$

Where C represents the total number of classes involved in the classification problem. The term p_i denotes the probability or proportion of data instances belonging to class $i$at a particular node in the decision tree. This measure is used to quantify the level of impurity at a node, where lower Gini values indicate more homogeneous class distributions.

E. Evaluation

Test data for every 90:10 and 80:20 data sharing scenario will be used to assess the trained model's performance. The Confusion Matrix findings, which contrast the model's predictions with the actual labels, will be used to compute the evaluation metrics. To assess the model's performance, a confusion matrix is employed. The purpose of this matrix is to compute and evaluate the effectiveness of a classification model. In assessing performance using a confusion matrix, there are four main elements used to identify the results of model predictions, namely: accuracy, precision, recall, and f1-score [21]. A representation of the confusion matrix can be seen in the following table, which shows the relationship between the model's prediction results and the actual conditions in the data [22]. The confusion matrix is presented in Table 3.

TABLE 3
CONFUSION MATRIX

Aktual	**Predicted Condition**	
	+	-
+	True Positive (TP)	False Negative(FN)
-	False Positive(FP)	True Negative (TN)

The metrics to be analyzed are:

Accuracy is used to measure the total percentage of predictions correctly classified by the model as a whole [22]. This metric provides an overview of the success rate of the model in predicting the appropriate class across all test data. The calculation of the accuracy value is based on a comparison between the number of correct predictions and the sum of all the data tested, as formulated in Equation (2).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{2}$$

Precision is used to measure the level of reliability of the model in predicting positive classes, in particular in minimizing false positive prediction errors [1], [24]. This metric shows how large the proportion of positive predictions is correct compared to all predictions classified as positive. The calculation of the precision value is according to the proportion of True Positive to the total of True Positive and False Positive, as shown by Equation (3).

$$Precision = \frac{TP}{TP+FP} \tag{3}$$

Recall is used to measure the model's ability to detect all actual phishing cases, so it is critical to minimize the risk of undetected threats [1], [22], [23]. This metric shows how much of the model's successful phishing data is correctly identified by the model from all existing phishing cases. The calculation of the recall value is based on the ratio between the True Positive and the sum of True Positives and False Negatives, as formulated in Equation (4).

$$Recall = \frac{TP}{TP+FN} \tag{4}$$

The F1-Score is used to provide a balanced measure of model performance by combining precision and recall values in a single metric [22], [23]. Because it illustrates the trade-off between the model's capacity to prevent false positive mistakes and its inability to identify positive situations, this statistic is especially helpful when there is a class imbalance. The harmonic mean of accuracy and recall is used to calculate the F1-Score, as formulated in Equation (5).

$$F1-Score = 2 . \frac{Precision\ .\ Recall}{Precision+Recall} \tag{5}$$

The results of all test scenarios will be thoroughly analyzed and compared to evaluate the effect of differences in data sharing ratios on model performance. This comparison is made based on the values of the evaluation metrics used, namely accuracy, precision, recall, and F1-score. Through this analysis, it can be determined the ratio of training data sharing and test data that produces the most optimal and stable model performance.

IV. Results

The findings of an experiment to assess how well the Random Forest algorithm classifies phishing URLs using a variety of pre-established test situations are shown in this section. Two data sharing ratios, 80:20 and 90:10, were used for the evaluation, with differences in the parameters of the number of trees (n_estimators) and the maximum depth of trees (max_depth). Furthermore, two feature configurations were tested: RF Full, which used all features, and RF Top-30, which used the top 30 characteristics chosen according to feature relevance. Table 4 summarizes the test results for all situations and compares the accuracy, precision, recall, and F1-score values for each combination of parameters and data partition ratio.

TABLE 4
COMPARISON OF TEST RESULTS IN EACH SCENARIO

Method	RASIO	n_estimator	Max_depth	Accuracy	Precision	Recall	F1-Score
RF Full	80:20	100	10	99.98	99.96	100	99.98
RF Full	80:20	100	20	99.99	99.99	100	99.99
RF Full	80:20	100	None	100	100	100	100
RF Full	80:20	200	10	99.99	99.98	100	99.99
RF Full	80:20	200	20	99.99	99.99	100	99.99
RF Full	80:20	200	None	100	100	100	100
RF Full	90:10	100	10	99.99	99.98	100	99.99
RF Full	90:10	100	20	100	100	100	100
RF Full	90:10	100	None	100	100	100	100
RF Full	90:10	200	10	99.99	99.99	100	99.99
RF Full	90:10	200	20	100	100	100	100
RF Full	90:10	200	None	100	100	100	100
RF Top-30	80:20	100	10	100	100	100	100
RF Top-30	80:20	100	20	100	100	100	100
RF Top-30	80:20	100	None	100	100	100	100
RF Top-30	80:20	200	10	100	100	100	100
RF Top-30	80:20	200	20	100	100	100	100

RF Top-30	80:20	200	None	100	100	100	100
RF Top-30	90:10	100	10	100	100	100	100
RF Top-30	90:10	100	20	100	100	100	100
RF Top-30	90:10	100	None	100	100	100	100
RF Top-30	90:10	200	10	100	100	100	100
RF Top-30	90:10	200	20	100	100	100	100
RF Top-30	90:10	200	None	100	100	100	100

It is evident from Table 4 that the Random Forest model performs exceptionally well and consistently in every test situation. In both the 80:20 and 90:10 data sharing ratios, the accuracy, precision, recall, and F1-score values fall between 99.96% and 100%. These findings show that the model has outstanding generalization skills to test data and can distinguish between phishing and legal URLs with a very low error rate.

In the RF Full configuration, an increase in max_depth values from 10 to 20 and None tends to result in improved performance, with some scenarios achieving perfect values across evaluation metrics. This suggests that overly tight tree depth restrictions can limit the model's ability to capture the complexity of data patterns, while max_depth = None provides greater flexibility without degrading performance. Meanwhile, variations in the number of trees (n_estimators) between 100 and 200 showed no significant difference in performance, so the use of a lower number of trees could be chosen to improve computing efficiency without sacrificing model accuracy.

The results of the test on the RF Top-30 configuration, as shown in Table 3, show that the use of only the top 30 features of the feature importance selection is capable of producing performance equivalent to a model that uses all features. In all test scenarios, RF Top-30 consistently achieved 100% scores on the accuracy, precision, recall, and F1-score metrics, indicating that the feature selection process did not degrade the model's classification capabilities. Comparisons between RF Full and RF Top-30 indicate that most of the predictive information in the dataset is concentrated on a small number of the most significant features, so that the complexity of the model can be reduced without sacrificing classification performance.

To provide a more detailed picture of the contribution of each feature, Figure 3 shows a visualization of the top 30 features based on feature importance values in one of the test scenarios, namely the 80:20 data sharing ratio with n_estimators = 100 and max_depth = 10. The visualization shows that some features have a much more dominant contribution than others in the model's decision-making process. The dominance of these features indicates that certain characteristics of URLs have an important role in distinguishing between phishing URLs and legitimate URLs.

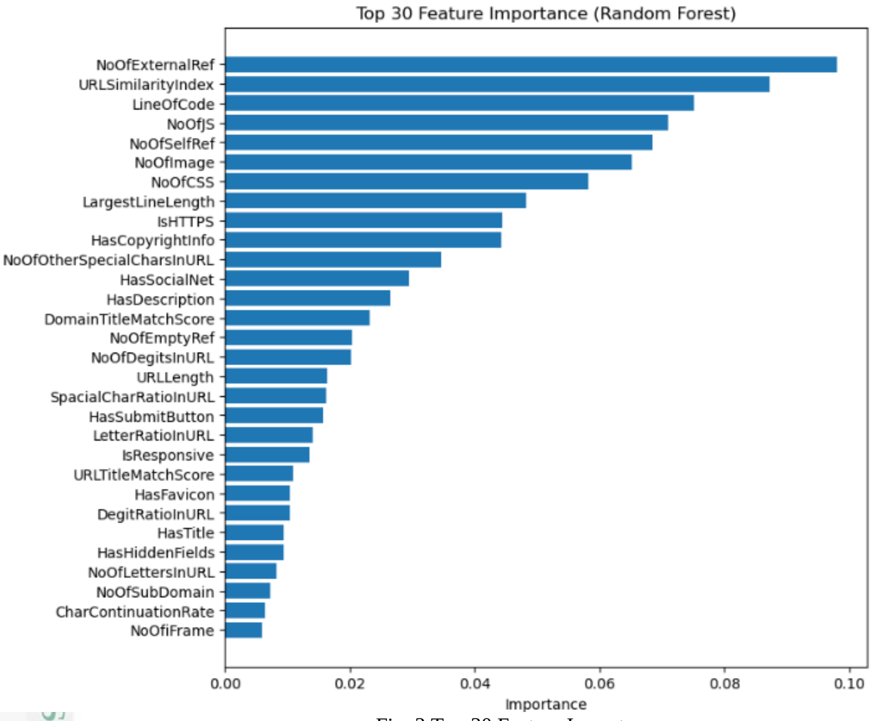


Fig. 3 Top 30 Feature Importance

Figure 3 illustrates how some characteristics have a significant impact on the Random Forest model's decision-making process. These criteria often have to do with the URL's structural aspects, such its length, the inclusion of special characters, and other technological characteristics that are frequently seen in phishing URLs. The dominance of these features suggests that certain patterns in the URL structure have a high discriminating power in distinguishing between phishing URLs and legitimate URLs. This condition also explains why the model is still able to achieve very high classification performance even though the number of features is significantly reduced through the feature selection process.

Overall, the results of the experiment showed that the Random Forest algorithm was able to provide very high and stable classification performance on various parameter configurations and data partition ratios. The application of feature importance-based feature selection has proven to be effective in simplifying the model without degrading performance, and is even able to maintain accuracy, precision, recall, and F1-score values at the maximum level. With lower model complexity and better computing efficiency, RF Top-30 configurations can be considered as a more optimal alternative to RF Full, especially in the development of URL phishing detection systems that require high accuracy, model stability, and computational resource efficiency.

V. DISCUSSION

The results of the experiment showed that the Random Forest algorithm produced very high and stable classification performance in detecting phishing URLs across various parameter configurations and data sharing ratios. The consistency of accuracy, precision, recall, and F1-score values that are close to or reaching 100% indicates that the model is able to effectively learn the characteristic patterns of phishing URLs. These findings are in line with previous research that confirmed Random Forest's superiority in handling high-dimensional data as well as URL-structure-based features in phishing [4], [10]. In addition, these results are in line with the findings of Ibrahim et al. who showed that ensemble-based models have stable and robust performance against parameter variations, making them effective in detecting complex cybersecurity threats compared to single models [24].

One of the main contributions of this research is the successful implementation of feature importance-based feature selection. Using only the top 30 features has been proven to maintain classification performance on par with the use of all features, while significantly lowering the complexity of the model. These findings reinforce the results of previous research that stated that the right selection of features can improve computational efficiency and maintain model accuracy in URL-based phishing detection [15]. In addition, this approach is also consistent with recent research proposing a URL-based phishing detection system using a hybrid machine learning approach, where the selection of relevant features is a key factor in increasing detection effectiveness without excessively increasing computational complexity [27]

Although the results obtained are very promising, this study has some limitations. Model evaluation is only conducted using one dataset, so the model's generalization ability against other phishing datasets with different data distribution characteristics cannot be fully ascertained. In addition, the tests were conducted in an offline environment, so the model's performance in real-time scenarios, such as direct phishing detection on a network system or web application, has not been thoroughly evaluated. This limitation needs to be considered in interpreting the results of the research.

As a further direction of research, it is recommended to conduct cross-dataset evaluation to test the robustness and generalization capabilities of the model against different types of phishing URLs. In addition, testing in a real-time environment as well as exploring hybrid approaches that combine Random Forest with other machine learning algorithms or deep learning techniques can be promising developments. The integration of the model into a web-based phishing detection system can also be a follow-up step to test the effectiveness of the model under real operational conditions. This study also complements the findings of Rawla et al. [15]. on similar datasets. In contrast to their results where *Random Forest's* performance (98.69%) was below *Deep Learning* (99.38%), this study shows a different perspective. Through *preprocessing* optimization and feature selection (*RF Top-30*), *Random Forest* is proven to be able to achieve maximum accuracy (100%). This indicates that a more compact model with 30 key features is already quite adequate to recognize attack patterns on this dataset without requiring high computational complexity.

VI. CONCLUSIONS

This study explicitly answers the research question regarding the effectiveness of the Random Forest algorithm in detecting phishing URLs in large-scale modern datasets. According to the findings of the trials, Random Forest is capable of producing very high and steady classification performance, with accuracy, precision, recall, and F1-score values near to or above 100% on various parameter configurations and data sharing ratios. In addition, test results showed that the variation in n_estimators and max_depth parameters did not have a significant effect on performance improvement after the model reached the optimal configuration. These results validate Random Forest as an efficient and dependable method for URL-based phishing detection, especially when considering model stability and efficiency. From a practical perspective, these findings indicate that Random Forest-based phishing detection systems can be reliably deployed without extensive parameter tuning, making them particularly suitable for organizations with limited computational resources or technical expertise.

The use of feature importance-based feature selection, which can simplify the model without compromising classification performance, is the study's primary contribution. Using only the top 30 features has been shown to deliver results equivalent to using the entire feature, while lowering the complexity and computational needs of the model. Thus, this study makes a practical contribution in the form of recommendations for the configuration of the Random Forest model that is efficient and easy to implement for the URL phishing detection system. For future work, it is recommended that this model be evaluated across different phishing datasets to assess its generalization capability, as well as tested in real-time detection scenarios to measure its effectiveness under operational conditions. It is anticipated that these discoveries will provide the foundation for the creation of cybersecurity systems that are precise, effective, and prepared for implementation in an actual operational setting.

REFERENCE

[1] A. F. Mahmud and S. Wirawan, “Sistemasi: Jurnal Sistem Informasi Deteksi Phishing Website menggunakan Machine Learning Metode Klasifikasi Phishing Website Detection using Machine Learning Classification Method,” vol. 13, no. 4, pp. 2540–9719, 2024, [Online]. Available: http://sistemasi.ftik.unisi.ac.id

[2] L. Ikhwanul Uzlah, R. Adi Saputra, and I. Isnawaty, “Deteksi Serangan Siber Pada Jaringan Komputer Menggunakan Metode Random Forest,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 3, pp. 2787–2793, 2024, doi: 10.36040/jati.v8i3.8891.

[3] A. K. Kencana, F. D. Ananda, A. D. Hartanto, and H. Hartatik, “Implementasi Metode Random Forest Klasifikasi untuk Phishing Link Detection,” *Intechno J. (Information Technol. Journal)*, vol. 4, no. 2, pp. 55–59, 2022, doi: 10.24076/intechnojournal.2022v4i2.1562.

[4] W. Sarasjati, S. Rustad, Purwanto, H. A. Santoso, and D. R. I. M. Setiadi, “Phishing Detection Using Random Forest-Based Weighted Bootstrap Sampling and LASSO+ Feature Selection,” *Int. J. Saf. Secur. Eng.*, vol. 14, no. 6, pp. 1783–1794, 2024, doi: 10.18280/ijsse.140613.

[5] T. S. Siddhesh, S. M. Rajagopal, and S. Bhaskaran, *Comparative Analysis of Machine Learning Algorithms for Anomaly Detection*, no. March. Springer Singapore, 2024. doi: 10.1109/I2CT61223.2024.10544217.

[6] R. Yang, K. Zheng, B. Wu, C. Wu, and X. Wang, “Phishing website detection based on deep convolutional neural network and random forest ensemble learning,” *Sensors*, vol. 21, no. 24, pp. 1–18, 2021, doi: 10.3390/s21248281.

[7] M. A. G. Al Ghifani, B. Hananto, and B. T. Wahyono, “Implementasi Ekstensi Google Chrome Dalam Mendeteksi Situs Web Phishing Menggunakan Algoritma Random Forest,” *Semin. Nas. Mhs. Ilmu Komput. dan Apl.*, vol. 3, pp. 179–189, 2022, [Online]. Available: https://conference.upnvj.ac.id/index.php/senamika/article/view/2152%0Ahttps://conference.upnvj.ac.id/index.php/senamika/article/download/2152/1744

[8] Y. S. Murti and P. Naveen, “Machine Learning Algorithms for Phishing Email Detection,” *J. Logist. Informatics Serv. Sci.*, vol. 10, no. 2, pp. 249–261, 2023, doi: 10.33168/JLISS.2023.0217.

[9] V. A. Windarni, A. F. Nugraha, S. T. A. Ramadhani, D. A. Istiqomah, F. M. Puri, and A. Setiawan, “Deteksi Website Phishing Menggunakan Teknik Filter Pada Model Machine Learning,” *Inf. Syst. J.*, vol. 6, no. 01, pp. 69–80, 2023, doi: 10.24076/infosjournal.2023v6i01.1268.

[10] T. O. Ojewumi, G. O. Ogunleye, B. O. Oguntunde, O. Folorunsho, S. G. Fashoto, and N. Ogbu, “Performance evaluation of machine learning tools for detection of phishing attacks on web pages,” *Sci. African*, vol. 16, p. e01165, 2022, doi: 10.1016/j.sciaf.2022.e01165.

[11] N. B. Putri and A. W. Wijayanto, “Analisis Komparasi Algoritma Klasifikasi Data Mining Dalam Klasifikasi Website Phishing,” *Komputika J. Sist. Komput.*, vol. 11, no. 1, pp. 59–66, 2022, doi: 10.34010/komputika.v11i1.4350.

[12] D. Sarma, T. Mittra, R. M. Bawm, T. Sarwar, F. F. Lima, and S. Hossain, “Comparative Analysis of Machine Learning Algorithms for Phishing Website Detection,” *Inventive Computation and Information Technologies*, pp. 883–896, 2021, doi: https://doi.org/10.1007/978-981-33-4305-4_64.

[13] Ary Prandika Siregar, Dwi Priyadi Purba, Jojor Putri Pasaribu, and Khairul Reza Bakara, “Implementasi Algoritma Random Forest Dalam Klasifikasi Diagnosis Penyakit Stroke,” *J. Penelit. Rumpun Ilmu Tek.*, vol. 2, no. 4, pp. 155–164, 2023, doi: 10.55606/juprit.v2i4.3039.

[14] A. Shah, “Classification and Detection of email Phishing using random Forest supervised-unsupervised machine learning algorithms,” 2021, [Online]. Available: http://norma.ncirl.ie/id/eprint/5126%0Ahttp://norma.ncirl.ie/5126/1/akshatshah.pdf

[15] A. Rawla, S. Singh, and P. Dubey, “ Detection of Phishing Attacks in PhiUSIIL Dataset using Deep Detection of Phishing Attacks in PhiUSIIL Dataset using Deep Learning Learning,” *Procedia Comput. Sci.*, vol. 259, pp. 543–552, 2025, doi: 10.1016/j.procs.2025.04.003.

[16] A. A. Akinyelu and A. O. Adewumi, “Classification of Phishing Email Using Random Forest Machine Learning Technique,” *Journal of Applied Mathematics*, vol. 2014, pp. 1–6, 2014, doi:

https://doi.org/10.1155/2014/425731.

[17] J. J. M. L. Piñeiro and L. R. Wong Portillo, "Web architecture for URL-based phishing detection based on Random Forest, Classification Trees, and Support Vector Machine," *Intel. Artif.*, vol. 25, no. 69, pp. 107–121, 2022, doi: 10.4114/intartif.vol25iss69pp107-121.

[18] M. G. HR, A. MV, G. P. S, and V. S, "Development of anti-phishing browser based on random forest and rule of extraction framework," *Cybersecurity*, vol. 3, no. 1, Oct. 2020, doi: https://doi.org/10.1186/s42400-020-00059-1.

[19] M. A. V. Ideal and I. Fitriyanto, "Implementation of a K-Means-Based Intelligent Patient Complaint Clustering System to Identify Handling Priorities," *Knowbase Int. J. Knowl. Database*, vol. 5, no. 1, pp. 69–80, 2025, doi: 10.30983/knowbase.v5i1.9529.

[20] S. Chandra and A. Prasad, "PhiUSIIL: A diverse security profile empowered phishing URL detection framework based on similarity index and incremental learning," *Comput. Secur.*, vol. 136, 2024, doi: https://doi.org/10.1016/j.cose.2023.103545.

[21] C. Umam, L. B. Handoko, and F. O. Isinkaye, "Performance Analysis of Support Vector Classification and Random Forest in Phishing Email Classification," *Sci. J. Informatics*, vol. 11, no. 2, pp. 367–374, 2024, doi: 10.15294/sji.v11i2.3301.

[22] A. A. Ubing, S. Kamilia, B. Jasmi, A. Abdullah, N. Z. Jhanjhi, and M. Supramaniam, "Phishing Website Detection : An Improved Accuracy through Feature Selection and Ensemble Learning," vol. 10, no. 1, pp. 252–257, 2019.

[23] H. Hidayat, A. Sunyoto, and H. Al Fatta, "Klasifikasi Penyakit Jantung Menggunakan Random Forest Clasifier," *J. SISKOM-KB (Sistem Komput. dan Kecerdasan Buatan)*, vol. 7, no. 1, pp. 31–40, 2023, doi: 10.47970/siskom-kb.v7i1.464.

[24] A. Hanafi *et al.*, "Implementasi Algoritma Random Forest Untuk Mendeteksi Penyakit Multiple Sclerosis," *National Conference on Electrical, Informatics and Industrial Technology (NEIIT)*, Vol. 1, No. 1, 2024, [Online]. Available: http://ojs.ft.uniska-kediri.ac.id/index.php/neiit/article/view/93/24.

[25] M. Azhima, I. Afrianty, E. Budianita, and S. Kurnia Gusti, "KLIK: Kajian Ilmiah Informatika dan Komputer Penerapan Metode Backpropagation Neural Network untuk Klasifikasi Penyakit Stroke," *Media Online)*, vol. 4, no. 6, pp. 3013–3021, 2024, doi: 10.30865/klik.v4i6.1956.

[26] N. Ibrahim, N. R. Rajalakshmi, V. Sivakumar, and L. Sharmila, "An optimized hybrid ensemble machine learning model combining multiple classifiers for detecting advanced persistent threats in networks," *J. Big Data*, vol. 12, no. 212, pp. 1–28, 2025, doi: 10.1186/s40537-025-01272-w.

[27] A. Karim, M. Shahroz, K. Mustofa, S. B. Belhaouari, and S. R. K. Joga, "Phishing Detection System Through Hybrid Machine Learning Based on URL," *IEEE Access*, vol. 11, no. March, pp. 36805–36822, 2023, doi: 10.1109/ACCESS.2023.3252366.

KOMUNITAS
DOSEN INDONESIA

Editor Decision
KDI/1033/LoA-Invoice/2025
30/12/2025

Dears Mr/Mrs/Ms. **Melyana Hasibuan , Rahmad Abdillah, Surya Agustian, Reski Mai Candra**
Thank you for your trust in our services.

With this confirmation, we have decided on your submission to **bit-Tech ISSN. 2622-271X (Print) & 2622-2728 (Online)**, **Classification of Phishing URL Attacks Using Random Forest Algorithm Based on Feature Importance** has been Accepted, for Vol 8 No 2 according to the results of the reviewer's reading notes.

The article will be uploaded and published online with an estimated time is **10/12/2025** softcopy (e-version).

Thank you for your attention and cooperation.

Sincerely,

Komunitas Dosen Indonesia

payment at least 3 days after the article is received: 31/12/2025
Please confirm payment via wa: 081807834703
