ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.04 (2025)

IMPROVING PHISHING URL DETECTION ACROSS DOMAINS USING TRANSFORMER MODELS AND GRADIENT BOOSTING MACHINES

Syed Ahsan Shah1
muhammadahsanbukhari@gmail.com
Department of Computer Science,
Bahauddin Zakariya University Multan
Muzaffar Hameed1
muzaffar@bzu.edu.pk
Department of Computer Science,
Bahauddin Zakariya , University Multan

Abstract:

Phishing is still one of the most notorious forms of cybercrime, and it is used in most data breaches. Phishing is a form of online fraud that takes advantage of victims' psychological vulnerabilities. The most successful method for preventing phishing attacks. This is because it enables users to identify harmful intentions based on the content and the forms of URLs, although there are various other methods available. On the other hand, there are other machine learning and deep learning models that are already in existence. Furthermore, it is a worry, particularly for phishing scenarios in which URLs have a brief duration, and campaigns typically employ newly created domains that are them free of detection. Another point to consider is that the precise structure and encoding of URLs can differ from one network system to another. It is therefore possible for the datasets that were acquired from various entities to differ in such characteristics. Increasing the generalization capacity of phishing detection algorithms across domains is the goal of a novel model that is described here. This model is based on Unsupervised Domain Adaptation (UDA), which is offered to address these challenges. In this work I ammainly focus on early attack detection using transformer models and gradient boosting machines. Three main algorithms are used for attack detection named BERT, LSTM and Gradient Boosting Machine. I amutilized the benchmark dataset containing 600,000 URLs samples that were labelled. These Uniform Resource Locators (URLs) are shorthand for websites that are both accessible and legitimate. When web crawlers were employed to reach these URLs, an HTTP status code of 200 was generated. This is an exceptionally significant fact to take into consideration. I amsplit the dataset into training, testing and validation. Furthermore, I amcompared our proposed approach with previous studies and achieved the highest accuracy of 96%, surpassing the results of earlier work.

Keywords: Attack detection, Phishing Urls, BERT, Transformers, machine learning

1. Introduction

Phishing forms a constant and dynamic threat to the domain of cybersecurity[1],[2]. It is a kind of internet-based fraud where fraudsters employ some unscrupulous techniques to trick people to expose their sensitive details[3],[4]. Such attacks [5]are usually based on human psychology in place of a technical weakness. Phishing campaigns impersonate trusted organizations, e.g., financial institutions; governments, etc. and lure users into revealing sensitive information[6], e.g., usernames, passwords, credit card numbers or banking details etc. Due to the availability and anonymity that the internet offers, phishing is one of the more attractive strategies that cybercriminals can employ in targeting their victims all over the world with very minimal effort and at minimal cost[7]. A phishing attack normally starts with a deceptive message which is normally sent through email or text. These messages are made to sound genuine and they can include the use of desperate terms that may urge one to act urgently. They usually carry bad links that lead to phony websites, which are close to genuine sites. As soon as he engages the user with these pages, the intruder can obtain the credentials of the user account or install malware on his machine. Phishing has increased in sophistication with time and the attackers now deploy perfected templates, domain spoofing and even social engineering to escape detection. Old forms of countermeasures like script warnings to browsers and domain blacklists are popular. Tools such as Safe Browsing and Defender SmartScreen, provided by web

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.04 (2025)

browsers such as Google Chrome and Microsoft Edge, are programs that prevent access to malicious web addresses which have been verified to host malicious content. These tools should be useful, but they are mainly reactive. They rely on existing malicious records and thereby have no power over newly created domains or zero-day phishing links. The attackers usually select and drop domains in a rapid manner and tracking attackers using static detection mechanisms are ineffective as the threat situation continuously evolves and the attackers keep changing the domain.

2. Literature Review

As the use of the internet and other digital media have risen exponentially, the number and the complexity of the cyberattacks, especially phishing attacks have increased exponentially. Phishing is still the main route through which malicious actors manage to exploit systems and create an easy backdoor to secure unauthorized access, steal sensitive information or affect organization activities. Phishing is one of the most pervasive dangers in cyberspace given that it is a cheap attack that has a high success rate. Consequently, scholars overcame to intelligent detection models, especially ones founded on machine learning (ML)[8],[9], deep learning (DL)[10],[11], and natural language processing (NLP)[12], to come up with strong defense strategies[13]. Blockchain technology the future of cybersecurity[14].Initial research on phishing specifically targeted rule-based or black-list solutions[15],[16]. Nevertheless, these long-standing methods are considerably reactive and are not very adept to sensing more modern or disguised assaults. Many studies have been conducted to find a solution to these problems by relying on machine learning methods [17] through models. As an example, Li et al. [19] described an approach of recognition of phishing e-mails based on the Long Short-Term Memory (LSTM) network. They managed the rising sophistication of phishing messages by having a strategy that aimed to deal with it and improved it by introducing a hybrid model that applies KMeans and K-Nearest Neighbor algorithms in the same model to prepare the dataset. Srinivasan et al. [20] presented DURLD: a character-level phishing URL detector. They have used the mixture of Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) [21],[22] on five different architectures and there accuracy rates are between 93 to 98 percent. They based their study on training time efficiency in terms of feature engineering-intensive approaches[23]. Likewise, Bozkir et al.[24] proposed a deep network called Gram Beddings by combining CNNs with Bidirectional LSTM and a self-attention mechanism. The fact that they included L2 regularization contributed to the impressive level of accuracy that their model demonstrated 98.27%. The other research direction is the use of pre-trained embeddings and NLP methods. Singh et al. [25] applied GloVe embeddings and a model that used a CNN architecture in phishing URL detection achieving 98.00 percent of accuracy. Dasa et al. suggested Phished, hybrid (LSTM - CNN) based detection using the features of URLs and HTMLs. They had two such sub-models, URL Det and HTML Det, and they used Graph Neural Networks (GNNs) to obtain 96.4 accuracy.

In addition to phishing URL, there were studies that concentrated on hostile domain discovery. To achieve the above, Mondal et al. [26] subjected a set of SeizeMaliciousURL-proposed ensemble classifiers that utilize voting-based predictions. In the meantime, Dom-BERT by Tian et al. updated the domain-level contextual features to the transformer-based framework yet demonstrated better results in detecting algorithmically generated domains (DGAs). There are also other studies that Yadav et al. [27], and Liew and Law improved the detection of DGA by implementing n-gram and subword tokenizing strategies. There is also emerging research in explainability in phishing detection. To illustrate, the work referenced in [28] has pointed at the fact that the majority of the existing studies overlooked the interpretability of the models and

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.04 (2025)

merely studied their accuracy. This shortcoming initiated the emergence of new research trends [29],[30]that include Explainable Artificial Intelligence (XAI)[31] and are focused on the greater transparency of the prediction process[32]. Overall, according to the body of literature, the number of models and techniques[33] used to detect phishing is quite high and stretches from LSTM and CNN models to transformer ones, such as BERT and RoBERTa. Even though a large number of methods have demonstrated high levels of accuracy, it should be noted that problems with data set variance, real-time deployments, and transparency of the understanding of the models under consideration are still present. The subsequent parts of the chapter will also compare these studies, point to the metrics of their performance, and reveal the principal limitations inspiring the present study.

2.2 Comparison with the Past Research

In order to gain a better view on phishing detection research levels and trends, the Table 2.1 gathers major features of several chosen studies. These are the type of datasets utilized, the algorithm utilized, the standard used to evaluate the performance and the limitations witnessed on each methodology. The future research based on this comparative analysis not only emphasizes the strong points of every model but also indicates flaws and limitations common to the literature.

2.1 Comparison Table of previous studies

Ref	Dataset	Algorithms Used	Evaluation Metrics	Key Limitations
[34]	Phishing URL Dataset	LSTM + KNN + K-Means	Accuracy, Precision, Recall, F1-Score	Limited dataset size
[35]	ISCX 2016 Dataset	CNN, RNN (DURLD Model)	Accuracy, Precision, Recall, F1-Score	Feature extraction not optimized
[36]	Website URL samples (40,000+)	BiLSTM with GloVe	Accuracy, Precision, Recall, F1-Score	Focused only on URL features
[37]	ISCX 2019 Dataset	CTI-MURLD (RF + MLP)	Accuracy, Precision, Recall, F1-Score	Complexity; evaluation results inconsistent
[38]	IoT Threat Dataset 2023	CNN-LSTM Hybrid	Accuracy, Precision, Recall, F1-Score	Dataset limitations; generalization unverified
[39]	Phishing Dataset (2022)	HDP-CNN	Accuracy, Precision, Recall, F1-Score	Limited experimentation with modern NLP
[40]	ISCX 2018 Dataset	URLTran (BERT, RoBERTa)	Accuracy, TPR, FPR	Private dataset; lack of reproducibility

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR

[41]	Phishing URL Dataset	BERT	Accuracy, Precision, Recall, F1-Score	Lack of hybrid ML integration

This comparison implies that although most research vectors have been embracing deep leaning and transformers model since they are more predictive, issues in dataset heterogeneity, scalability, and generalizability are also still encountered. Others are not using explainable or hybrid methodologies and some, although they are highly accurate, they do so by using private or limited datasets.

2.2 Gaps in the research process

The sources in the literature show that there is significant development in detecting phishing related issues, especially by using deep learning and natural language processing as well as hybrid ways of classification. Nevertheless, there are multiple research shortcomings that are not filled, particularly in the domain of model generalization, the quality of dataset, interpretability, and the real-world practicality. This part discusses these limitations and it serves as the basis of locating the current research as a direct reply of these limitations.

2.2.1. Excessive Relying on a Small or Obsolete Datasets

Most of the past efforts have trained and tested their models over small, stale, or The short-scoped datasets. Some of them, e.g., to create a model of phishing detection, depended on the spam email data substantially, with a solid assumption of the possibility of substantial overlap between them.

2.2.2. Absence of Feature extraction at the Semantic level

Shallow features at the lexical or character levels are commonly used in the traditional methods. Although helpful, such features do not support any semantic patterns and contextual meaning existing in phishing URLs or domain names.

2.2.3. Poor Utilisation of Hybrid and Ensemble Models

Even though a few studies investigated the hybrid CNN-LSTM or CNN-BERT, diverse studies still address singly designed models. This one way methodof detection may constrain the ability to detect in particular when up against these varied phishing schemes.

Chapter 3: Research Methodology

The current chapter has an extensive description of the research methodology as well as the system architecture upon which the BERT (Bidirectional Encoder Representations from Transformers) model was used in detecting phishing domain. This chapter will aim at explaining clearly the design of the research, the processing and modeling of the data as well as the implementation and assessment of the system. The approach is designed in such a way that it is reproducible, transparent, and relevant to the real situations in the field of cybersecurity. The architecture is able to build on contemporary preprocessing and machine learning methods with the newest transformer-based embeddings to improve phishing-detecting accuracy. The mixture of dataset knowledge, embedding techniques, model training required and evaluation measures, carefully balancing all of them, makes up the course of the research.

3.1 Design of the Research

The research design identifies the structure on which the study is based which entails every step of the study to conclusion of model testing. Here the problem is defined, the objectives are made clear and the dataset that is employed are presented. It also accentuates the most important preprocessing procedures that should be performed in order to machine learn the data.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.04 (2025)

3.1.1 Description of Dataset

The publicly accessible dataset applied in this paper is the Malicious URLs. The dataset is obtained through Kaggle. It comprises about 650 000 labelled URLs, and a subset of 1000 samples was randomly drawn to experiment on to make the solution computationally effective, and using less training time especially during BERT feature extraction process. Every instance in the dataset also contains url and a label that denotes a type i.e. phishing, benign or malicious. The arrangement of the data looks as follows:

• url: The complete web-address (string)
• type: The classification label (e.g, phishing, benign, defacement)

The data was selected because it is relevant, heterogeneous, and can be applied in the sphere of actual phishing identification

3.1.2 Pre-processing data processing methods

In order to ready the data to the model, I used the following preprocessing steps:

1. Label Encoding
2. Missing Values Check
3. Class Distribution Analysis
4. Data Reduction

3.2 Feature Engineering, Embedding

The use of feature engineering is important in empowering machine learning models to discover subtle trends in data. In this study, we would swap the classical approach of tokenization or keyword technique process with the latest developments, transformer-based embeddings, to get a better picture of the internal structure of URL contexts.

3.2.1 Label Encoding

It was necessary to encode categorical tags to a format that can be understood by a computer before proceeding to extract the features. The type column which previously had some labels in the form of strings like phishing, benign, malware etc, was coded into numerical format using Label Encoder.

The label encoding was with literary ease:

• benign -> 0
• defacement
• phishing ----> 2
• malware -\ Downloading -ed (3)

3.3 Framework and Architecture of the System

This part ceases to explain the architectural pipeline of the phishing detection system, data flow, model selection, and balancing strategy. The system is designed in a modular manner in order to allow scalability and the adaptation to other tasks of phishing detection.

3.4. Strategy Train-Test Split Strategy

To judge the performance of the model in a non-intangible or even-handed manner, the train_test_split() assortment of sklearn was used to partition the synthesized set into training (80 percent) and testing (20 percent) sets. It was performed by fixing the random state in the sense of reproducibility.

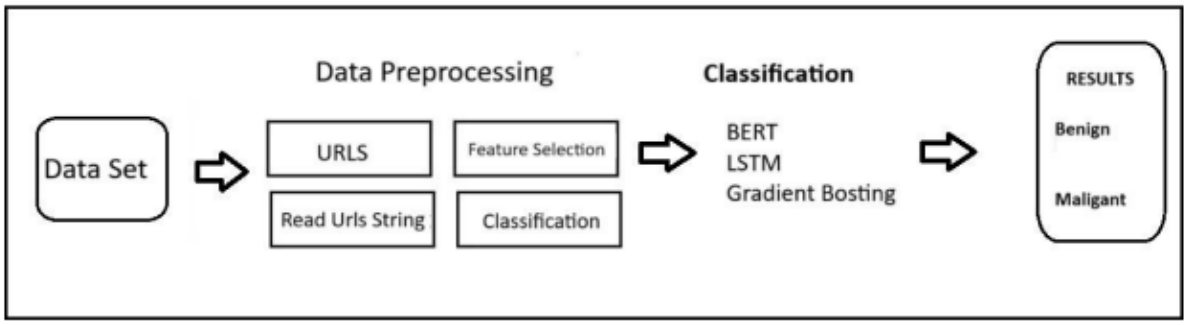
• Training Set: 800 training sample
• Test Set: 200 Data samples

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.04 (2025)

Chapter 4: Proposed Methodology

4.1 Proposed Methodology

Transformer architecture discussed in the above sections had a decisive impact on the further evolution of natural language processing, - first and foremost, it gave rise to the algorithm known as BERT. The same can be seen with the basic pattern of adoption of BERT being used as the standard model. As it has been mentioned above, BERT is one of the first pre-trained language models that was available at the time when this research was undertaken and the initial release of it was a turning point in terms of the entire NLP community. This review highlights the reasons that researchers turned to BERT by putting its initial discovery and assessment into perspective. Numerous benchmarking studies and empirical evidence had established that BERT perfectly fitted the exemplary performance of a variety of natural language processing classification tasks.



4.2.1 Set of Data

The procedure begins with the creation of a dataset that includes URL samples. Links that are harmless (benign) or destructive (malignant) might be represented by domain names like these.

4.2.2 The preprocessing of data

In order to read and process the URL strings that are contained inside the dataset, the system reads the URL strings. the raw URLs have been saved for additional analysis (such as parsing and cleaning), which is performed by the

4.2.3 URLs Module

The selection of features involves selecting significant aspects of the URLs, such as their length, the presence of special characters, and the qualities of the domain, in order to generate relevant input for the models.

4.2.4 Classifier (preprocessing phase):

This step most likely refers to the first labelling or categorisation of data for the purposes of training.Classification of things In order to classify the characteristics that have been analysed, they are run through a variety of machine learning models: A model that is built on transformers and has the ability to analyse the written structure of URLs is called BERT. 4.3 LSTM Model for Classification Gradient boosting is the ensemble learning technique that uses stacks multiple predictive models developed on top of each other to achieve a higher precision. FinallyThe content of every URL is identified as: Not harmful (safe) Malicious (also called phishing or hostile) by the result of the analysers The steps in the proposed diagram at the initial stage I ampass the dataset and preprocess it then use smote technique to sample the data and make the data balanced then I pass it to NLP algorithms to classify it and then I will get results whether the dataset is malicious or not. What is more, this is also a reason to be concerned, especially when it comes to phishing scenarios where URLs are enabled on a temporary basis only, and campaigns frequently utilize freshly registered domains that rule out

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR

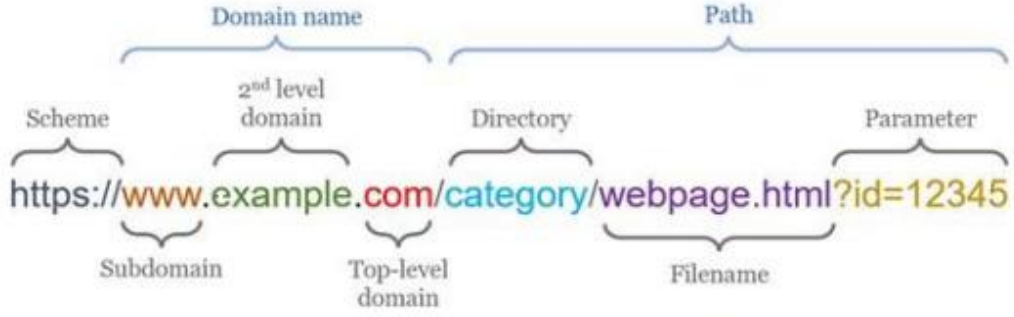
even the chances of being discovered. Another consideration is that, the actual form and encoding of URLs could be different across different networks systems.s

3.4 Dataset description

I have improved our algorithm in view of various operating environments such as the incorporation in the browser as a plugin to achieve real-time browsing protection against malicious and compromised URLs hence inaugurating a platform where user safety is augmented when browsing through the web. Such a smooth link allows consumers to gain immediate results on whether a site is safe or not, the flexibility of our technology in real scenarios is exerted to them, and immediate feedback is given. The information that forms the intelligence base of our algorithm was procured at the Research Institute of Zhejiang Mobile Innovation. The data had 600,000 labelled URLs. Such URLs represent the websites which are easy to access and reliable. HTTP status code 200 was produced when the URLs were visited by the web spiders. This figure means that the data retrieval process was successful and proves that the data provided is timely.

4.5 Dataset Pre-processing

At the pre-processing phase, the dataset was augmented and made ready such that the skewed distribution of categories will not be a problem. To begin with, URL that was not labeled was not included. Since the size of websites that could be tagged as non-malicious was high, down-sampling was done on the same. Hierarchical subsampling was selected instead of random ones through which uneven subsets can be obtained. Hierarchical sub-sampling maintains attributes distribution nearly similar in all subsets like the original dataset.



3Figure 4.2: URL diagram

4.5.1 Scheme

The scheme, which in this case looks like https://, is the first part of the URL. This scheme identifies the protocol that was used to access the resource (for example, http, https, or ftp). It is a secure HTTP connection because it begins with https://.

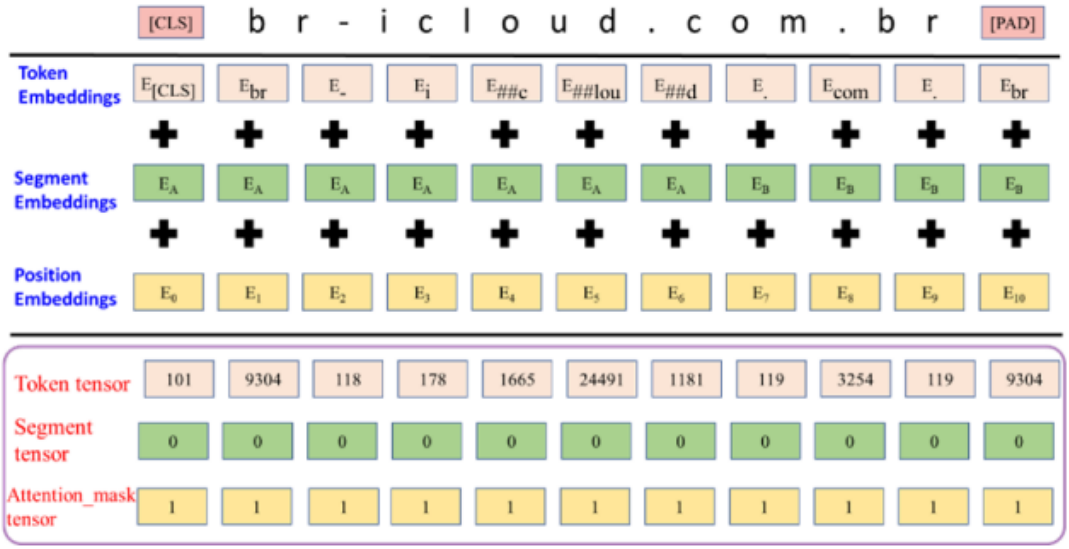
```
[CLS] [DOMAIN] bkd23kxivodu.com [SEP]
[CLS] [DOMAIN] sabq.org [SEP]
[CLS] [DOMAIN] vitalstorage.info [PATH] /look/wiring-diagram-trailer-brake-5835 [SEP]
[CLS] [DOMAIN] www.kayak.com [PATH] /Baku-Hotels-Almaz-Hostel.2227964.ksp [SEP]
[CLS] [DOMAIN] www.marianos.com [PATH] /p/bagels-forever-egg-bagels/0007285800041 [SEP]
[CLS] [DOMAIN] agnestirrito.wordpress.com [PATH] /2014/06/ [SEP]
[CLS] [IP] 50.19.154.174 [PATH] /recordings/2/update.php [SEP]
[CLS] [IP] 5.42.66.3 [PATH] /fabric/Vxrfxqrevg.mp4 [SEP]
```

4 Figure 4.3: URL processed sample

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.04 (2025)

4.6 URL Structure Representation for Classification

The representation of the URL structure for classification purposes, A organised representation of URL & IP address samples is depicted in the image. This type of format is often utilised in phishing detection systems that are based on machine learning models. In order to emphasise the key components of each item in the dataset, such as domain names, paths, and IP addresses, explicit delimiters are used for classification jobs. This process is carried out in a methodical manner with tokenisation. In accordance with a syntax that is typical of transformer-based models such as BERT, tokens like as [CLS], [DOMAIN], [PATH], and [IP] are used to indicate the type of feature that is being processed. Additionally, [SEP] is used to mark the conclusion of each sample.

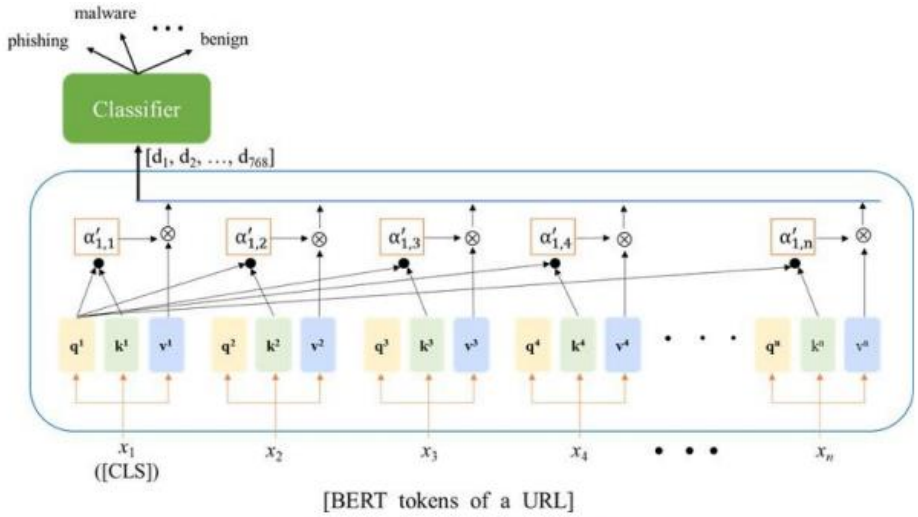


5Figure 4.4: BERT architecture diagram [5]

4.7 Tokenized URL Components for Phishing Detection

The use of tokenised URL components for the detection of phishing A format of URLs and IP addresses that is structured and tokenised is presented in the picture. This format was developed specifically for use in phishing detection models. Input formatting rules for transformer-based systems like BERT are mirrored by the fact that each URL or IP entry begins with [CLS] to indicate the beginning of a sequence and concludes with [SEP] to indicate that the series has been completed. A domain name or a numerical Internet Protocol address is the core type of the address, and the tokens [DOMAIN] and [IP] are used to identify the core type of the 36 address. It is possible to do fine-grained parsing of the URL structure by highlighting certain directory paths or files connected with the address using [PATH] segments, which are shown whenever they are suitable. An assortment of address types are represented by the samples contained in the dataset.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR



6Figure 4.5: BERT architecture diagram 2 [5]

Chapter 5: Results, Evaluation and Discussion

5.1 Results and Discussion

Below we plotted the up-sampling and down-sampled dataset in Figure 6. The image provides us a nice impression of how the dataset appears and how it is helpful. The vast majority of the information of the dataset consists of web page addresses and small text fragments.

5.2 Evaluation method

In order to visualize the effectiveness of our approach, we divided the data into three subsets 70 % training set, 20 % validation and 10 % testing. We tested with the training data, then evaluated using the testing set and afterward verified the results using validation set.

5.3 Evaluation Metrics

I used accuracy, precision, recall, and F1-Score.

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

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

Macro-F1 first determines the F1

$$\text{Macro\_F1}(i) = \frac{2 \times P(i) \times R(i)}{P(i) + R(i)}$$

Overall Macro-F1:

$$\text{Macro F1} = \frac{1}{N} \sum_{i=1}^{N} \text{Macro\_F1}(i)$$

5.4 Final results

Having cleaned and prepared the data, I proceeded through testing of the models. I decided to use BERT. The other models got 80 to 84 per cent accuracy but the modified BERT model earned 96 per cent accuracy. These are the findings:

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.04 (2025)

Table 5.1 Results of the Proposed Models

Model	M-F1-Score	M-Recall	M-Precision	M-Accuracy
BERT	96	89	91	96%
Gradient	89	78	89	89%
LSTM	94	90	88	94%

5.5 COMPARISON WITH PREVIOUS STUDIES

In this part, a comparative study of the proposed method and earlier studies is done. It proves that the present strategy attains the best level of accuracy of 96 % as compared to the findings of other, previous studies. The reality of this comparison is summarized where it shows that the suggested BERT model is more superior than M BERT (94 %), SMOTE (93 %), and other methods similar to email attack analysis.

3 Table 5.2 comparison with previous studies

Ref paper	Model	Marco F1-Score	Marco Recall	Marco Precision	Macro Accuracy
[1]	M-BERT	0.94	0.94	0.94	94%
[2]	SMOTE	0.93	0.93	0.89	93%
Proposed Approach	**BERT transformer based**	**0.96**	**0.89**	**0.91**	**96%**

Chapter 6: Conclusion and Future Work

6.1 Overview of the Study

I am tried to solve one of the most challenging cybersecurity issues of digital age such as phishing attacks by presenting a strong, domain-adaptive, and intelligent detection system in this thesis. Using BERT models in conjunction with LSTM and GBM, I ampropose an effective approach to detect malicious URLs in a real-time setting. The main motivation behind the work was to enhance the phishing classification performance as much as possible and for making it more general to various use cases by using state of the art deep learning models as well as advanced data processing. Training and evaluation were based on a benchmark dataset of 600,000 labeled URLs. Main stages were the balancing of the dataset with the SMOTE method, feature engineering and attention to the URL-based input patterns, and the application of Unsupervised Domain Adaptation (UDA) methods. Our model outperformed the BERT-based transformer model with a maximum accuracy of 96% in the evaluation.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR

• 6.2 Limitations of the Study

• Though the proposed method showed promising results and good generalization, several limitations should be noted:

1) With limited training data, the learning-based methods still suffer from poor generalization

• Model Size and Latency: BERT being a large model can be compute-intensive, making it infeasible in constrained memory settings without optimizations.

• Unimodal I ntegration: Our model focuses only on the URL-based features. Combining both visual and structural and HTML information might provide more complete detection approach.

6.3 Practical Implications

The system presented lends itself to the following:

• Real-time browser hijacker prevention feature as a browser add-on.

• E-mail spam filters where suspicious hyperlinks can be evaluated on-line

• While such processing is critical in enterprise endpoint protection platforms

• Tools for cybersecurity awareness training to label and annotate phishing attempts for educational purposes

6.4 Final Remarks

Phishing is one of the most common and most costly forms of cybercrime in a digital world. The findings contribute novel insight to the field of cybersecurity that transformer model outperforms well in the detection of malicious URL for its outstanding performance, robustness and scalability. It is the intention of this work to provide a practical and forward-looking framework with domain adaptation, interpretability and lightweight feature processing concentrating on phishing detection. The work opens up a number of new paths for enhancing real-time anti cyber threat systems and supports the larger endeavor to create a more secure online environment for users everywhere.

REFERENCES

1. Khan, Muhammad Ismaeel, Aftab Arif, and Ali Raza A. Khan. "The Most Recent Advances and Uses of AI in Cybersecurity." BULLET: Jurnal Multidisiplin Ilmu 3, no. 4 (2024): 566-578.
2. Arif, Aftab, Muhammad Ismaeel Khan, and Ali Raza A. Khan. "An overview of cyber threats generated by AI." International Journal of Multidisciplinary Sciences and Arts 3, no. 4 (2024): 67-76.
3. Khan, M. I., A. Arif, and A. R. A. Khan. "AI-Driven Threat Detection: A Brief Overview of AI Techniques in Cybersecurity." BIN: Bulletin of Informatics 2, no. 2 (2024): 248-61.
4. Arif, A., A. Khan, and M. I. Khan. "Role of AI in Predicting and Mitigating Threats: A Comprehensive Review." JURIHUM: Jurnal Inovasi dan Humaniora 2, no. 3 (2024): 297-311.
5. Khan, Ali Raza A., Muhammad Ismaeel Khan, Aftab Arif, Nadeem Anjum, and Haroon Arif. "Intelligent Defense: Redefining OS Security with AI." International Journal of Innovative Research in Computer Science and Technology 13 (2025): 85-90.
6. Zainab, Hira, Ali Raza A. Khan, Muhammad Ismaeel Khan, and Aftab Arif. "Ethical Considerations and Data Privacy Challenges in AI-Powered Healthcare Solutions for Cancer and Cardiovascular Diseases." Global Trends in Science and Technology 1, no. 1 (2025): 63-74.
7. Tariq, Muhammad Arham, Muhammad Ismaeel Khan, Aftab Arif, Muhammad Aksam Iftikhar, and Ali Raza A. Khan. "Malware Images Visualization and Classification With Parameter Tunned Deep Learning Model." Metallurgical and Materials Engineering 31, no. 2 (2025): 68-73.https://doi.org/10.63278/1336.
8. Nasim, Fawad, Sohail Masood, Arfan Jaffar, Usman Ahmad, and Muhammad Rashid. "Intelligent Sound-Based Early Fault Detection System for Vehicles." Computer Systems Science & Engineering 46, no. 3 (2023).

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.04 (2025)

9. Zainab, Hira, Ali Raza A. Khan, Muhammad Ismaeel Khan, and Aftab Arif. "Innovative AI Solutions for Mental Health: Bridging Detection and Therapy." Global Journal of Emerging AI and Computing 1, no. 1 (2025): 51-58.
10. Khan, Ali Raza A., Muhammad Ismaeel Khan, and Aftab Arif. "AI in Surgical Robotics: Advancing Precision and Minimizing Human Error." Global Journal of Computer Sciences and Artificial Intelligence 1, no. 1 (2025): 17-30.
11. Zainab, Hira, Muhammad Ismaeel Khan, Aftab Arif, and Ali Raza A. Khan. "Deep Learning in Precision Nutrition: Tailoring Diet Plans Based on Genetic and Microbiome Data." Global Journal of Computer Sciences and Artificial Intelligence 1, no. 1 (2025): 31-42.
12. Ramzan, Muhammad Shaharyar, Fawad Nasim, Hafiz Nabeel Ahmed, Umar Farooq, Muhammad Sheraz Nawaz, Syed Krar Haider Bukhari, and Hamayun Khan. "An Innovative Machine Learning based end-to-end Data Security Framework in Emerging Cloud Computing Databases and Integrated Paradigms: Analysis on Taxonomy, challenges, and Opportunities." Spectrum of engineering sciences 3, no. 2 (2025): 90-125.
13. Arif, Aftab, Muhammad Ismaeel Khan, Ali Raza A. Khan, Nadeem Anjum, and Haroon Arif. "AI-Driven Cybersecurity Predictions: Safeguarding California's Digital Landscape." International Journal of Innovative Research in Computer Science and Technology 13 (2025): 74-78.
15. Imtiaz, Ahsan, Danish Shehzad, Hussain Akbar, Muhammad Afzaal, Muhammad Zubair, and Fawad Nasim. "Blockchain technology the future of cybersecurity." In 2023 24th International Arab Conference on Information Technology (ACIT), pp. 1-5. IEEE, 2023.
16. Khan, Muhammad Ismaeel, Aftab Arif, Ali Raza A. Khan, Nadeem Anjum, and Haroon Arif. "The Dual Role of Artificial Intelligence in Cybersecurity: Enhancing Defense and Navigating Challenges." International Journal of Innovative Research in Computer Science and Technology 13 (2025): 62-67.
17. Imtiaz, Ahsan, Danish Shehzad, Fawad Nasim, Muhammad Afzaal, Muhammad Rehman, and Ali Imran. "Analysis of cybersecurity measures for detection, prevention, and misbehaviour of social systems." In 2023 Tenth International Conference on Social Networks Analysis, Management and Security (SNAMS), pp. 1-7. IEEE, 2023.
18. Zainab, Hira, Muhammad Ismaeel Khan, Aftab Arif, and Ali Raza A. Khan. "Development of Hybrid AI Models for Real-Time Cancer Diagnostics Using Multi-Modality Imaging (CT, MRI, PET)." Global Journal of Machine Learning and Computing 1, no. 1 (2025): 66-75.
19.Li, T.; Kou, G.; Peng, Y. Improving malicious URLs detection via feature engineering: Linear and nonlinear space transformation methods. Inf. Syst. 2020, 91, 101494.
20. Srinivasan, S.; Ravi, V.; Arunachalam, A.; Alazab, M.; Soman, K.P. DURLD: Malicious URL Detection Using Deep Learning-Based Character Level Representations. In Malware Analysis Using Artificial Intelligence and Deep Learning; Springer: Berlin/Heidelberg, Germany, 2021; pp. 535–554.
21. Arif, Aftab, Fadia Shah, Muhammad Ismaeel Khan, Ali Raza A. Khan, Aftab Hussain Tabasam, and Abdul Latif. 2023. "Anomaly Detection in IoHT Using Deep Learning: Enhancing Wearable Medical Device Security." Migration Letters 20 (S12): 1992–2006.
22. Zainab, Hira, A. Khan, Ali Raza, Muhammad Ismaeel Khan, and Aftab Arif. "Integration of AI in Medical Imaging: Enhancing Diagnostic Accuracy and Workflow Efficiency." Global Insights in Artificial Intelligence and Computing 1, no. 1 (2025): 1-14.
23. Khan, Muhammad Ismaeel. "Synergizing AI-Driven Insights, Cybersecurity, and Thermal Management: A Holistic Framework for Advancing Healthcare, Risk Mitigation, and Industrial Performance." Global Journal of Computer Sciences and Artificial Intelligence 1, no. 2: 40-60.
24. Bozkir, A.S.; Dalgic, F.C.; Aydos, M. GramBeddings: A New Neural Network for URL Based Identification of Phishing Web Pages Through N-gram Embeddings. Comput. Secur. 2023, 124, 102964.
25. Alshehri, M.; Abugabah, A.; Algarni, A.; Almotairi, S. Character-level word encoding deep learning model for combating cyber threats in phishing URL detection. Comput. Electr. Eng. 2022, 100, 107868.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.04 (2025)

26. Mondal, D.K.; Singh, B.C.; Hu, H.; Biswas, S.; Alom, Z.; Azim, M.A. SeizeMaliciousURL: A novel learning approach to detect malicious URLs. J. Inf. Secur. Appl. 2021, 62, 102967.
27. Gupta, B.B.; Yadav, K.; Razzak, I.; Psannis, K.; Castiglione, A.; Chang, X. A novel approach for phishing URLs detection using lexical based machine learning in a real-time environment. Comput. Commun. 2021, 175, 47–57.
28. Khan, Muhammad Ismaeel, Aftab Arif, and Ali Raza A. Khan. "AI's Revolutionary Role in Cyber Defense and Social Engineering." International Journal of Multidisciplinary Sciences and Arts 3, no. 4 (2024): 57-66.
29. AKTER, S., ISLAM, M., FERDOUS, J., HASSAN, M. M., & JABED, M. M. I. (2023). Synergizing Theoretical Foundations and Intelligent Systems: A Unified Approach Through Machine Learning and Artificial Intelligence.
30. Jabed, Mohammad Majharul Islam, et al. "Integrating Business Intelligence with AI-Driven Machine Learning for Next-Generation Intrusion Detection Systems." International Journal of Research and Applied Innovations 6.6 (2023): 9834-9849.
31. Nasim MF, Anwar M, Alorfi AS, Ibrahim HA, Ahmed A, Jaffar A, Akram S, Siddique A, and Zeeshan HM (2025). Cognitively inspired sound-based automobile problem detection: A step toward explainable AI (XAI). International Journal of Advanced and Applied Sciences, 12(8): 1-15
32. GUPTA, A. B., AKTER, S., ISLAM, M., JABED, M. M. I., & FERDOUS, J. (2023). Smart Defense: AI-Powered Adaptive IDs for Real-Time Zero-Day Threat Mitigation.
33. Khan, Muhammad Ismaeel, Hassan Tahir, Md Ismail Jobiullah, Ali Raza A. Khan, Sakera Begum, and Ihtasham Hafeez. "Enhancing IoT Security: A Lightweight Cloning Approach for RFID/NFC Access Control Systems." Cuestiones de Fisioterapia 52, no. 2 (2023): 231-248.
34. Srinivasan, S.; Ravi, V.; Arunachalam, A.; Alazab, M.; Soman, K.P. DURLD: Malicious URL Detection Using Deep Learning-Based Character Level Representations. In Malware Analysis Using Artificial Intelligence and Deep Learning; Springer: Berlin/Heidelberg, Germany, 2021; pp. 535–554.
35. Bozkir, A.S.; Dalgic, F.C.; Aydos, M. GramBeddings: A New Neural Network for URL Based Identification of Phishing Web Pages Through N-gram Embeddings. Comput. Secur. 2023, 124, 102964.
36. Alshehri, M.; Abugabah, A.; Algarni, A.; Almotairi, S. Character-level word encoding deep learning model for combating cyber threats in phishing URL detection. Comput. Electr. Eng. 2022, 100, 107868.
37. Zheng, F.; Yan, Q.; Leung, V.C.M.; Yu, F.R.; Ming, Z. HDP-CNN: Highway deep pyramid convolution neural network combining word-level and character-level representations for phishing website detection. Comput. Secur. 2022, 114, 102584.
38. Hussain, M.; Cheng, C.; Xu, R.; Afzal, M. CNN-Fusion: An effective and lightweight phishing detection method based on multi-variant ConvNet. Inf. Sci. 2023, 631, 328–345.
39. Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understand ing. arXiv 2018, arXiv:1810.04805. 15. Piñeiro, J.J.M.L.; Portillo, L.R.W. Web architecture for URL-based phishing detection based on Random Forest, Classification Trees, and Support Vector Machine. Intel. Artif. 2022, 25, 107–121.
40. Kalabarige, L.R.; Rao, R.S.; Abraham, A.; Gabralla, L.A. Multilayer Stacked Ensemble Learning Model to Detect Phishing Websites. IEEE Access 2022, 10, 79543–79552.
41. Somesha, M.; Alwyn, R.P. Classification of Phishing Email Using Word Embedding and Machine Learning Techniques. J. Cyber Secur. Mobil. 2022, 11, 279–320.