iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** III **Month of publication:** March 2024

DOI: https://doi.org/10.22214/ijraset.2024.59261

www.ijraset.com

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

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue III Mar 2024- Available at www.ijraset.com

Phishing URL Detection using Machine Learning

Sri Hari Nallamala[1], Kommu Namitha[2], Kunchanapalli Raviteja[3], Kadiyam Sai Sumanth[4], Jyothi Sri Kota[5]
[1]*Associate Professor,* [2,3,4,5] *UG Students, Department of CSE, Vasireddy Venkatadri Institute of Technology (Autonomous), Guntur, AP*

Abstract: ***Phishing attacks pose a significant threat in cyberspace, Cybercrimes such as Phishing means accessing personal data and violating security through the Internet and its main aim is to steal the information from the users using different techniques in that the primary one is Phishing which demands effective detection mechanisms.***
This study evaluates the performance of Gradient Boosting Classifier, Random Forest, and Decision Tree machine learning models in conjunction with feature selection techniques namely SelectKBest and Chi-Square. Initially, a comprehensive feature set of 30 attributes achieved a baseline accuracy of 97.4%. Through SelectKBest and Chi-Square feature selection methods, 13 key features were identified, leading to a slightly decreased accuracy of 95.6% upon model retraining. This research highlights the importance of feature selection in enhancing phishing detection accuracy while maintaining model interpretability. Technologies such as NumPy, Pandas, Matplotlib, Scikit-learn, and Flask drive.
The project, emphasising the exploration of ML models, EDA (Exploratory Data Analysis) on phishing data, and understanding feature importance. The Machine Learning Models like Gradient Boosting Classifier, Random Forest and Decision tree are used to detect whether the given URL is Malicious URL or Legitimate URL using Comprehensive Feature set and Key Feature set. This Models and Feature sets are compared based on the Performance Metrics namely Accuracy, Precision, Recall and F1-score.
This study contributes to advancing cyber defence mechanism through the fusion of sophisticated Machine Learning algorithms and meticulous Feature Selection methodologies.
Keywords: Machine Learning, Cyber Security, Gradient Boosting Classifier, Random Forest, Decision tree, EDA, Feature Selection, SelectKBest, Chi-Square.

I. INTRODUCTION

In today's digital landscape, the proliferation of cyber threats poses a significant risk to individuals and organisations worldwide. Among these threats, phishing is an insidious practice where malicious actors craft deceptive URLs to steal sensitive information. To combat this growing menace, the integration of Machine Learning (ML) techniques has emerged as a potent solution, empowering security systems to analyse vast datasets and detect patterns indicative of malicious activity.

This article is dedicated to developing a sophisticated phishing URL detection system, leveraging a diverse array of ML models such as the Gradient Boosting Classifier, Random Forest, and Decision Tree.

By considering a comprehensive feature set of 30 attributes and comparing the result with the 13 key features extracted from the 30 using feature selection techniques like SelectKBest and chi-Square, the system aims to classify these URLs as legitimate or malicious effectively.

With the ultimate goal of providing real-time protection against phishing attacks, the system can be seamlessly integrated into web browsers or security applications.

By harnessing technologies like NumPy, Pandas, Matplotlib, Scikit-learn, and Flask, the project endeavours to explore the nuanced intricacies of ML models, conduct Exploratory Data Analysis (EDA) on phishing datasets, and unravel the critical importance of various features in distinguishing between safe and malicious URLs.

This introduction sets the stage for a comprehensive exploration of how ML, when judiciously applied, can fortify cybersecurity measures, mitigate data breaches and financial loss risks, and safeguard users against the ever-evolving landscape of cyber threats.

©IJRASET: All Rights are Reserved

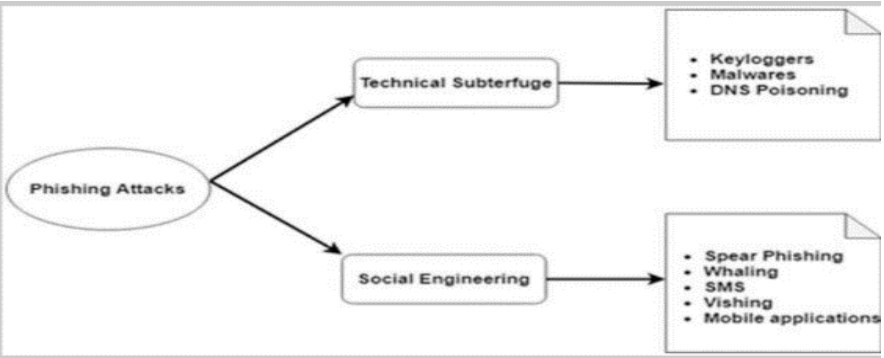


Fig.1: Multiple forms of phishing attacks.

II. LITERATURE REVIEW

Detecting fake sites has become an urgent task as we face increasing cyber-attacks that aim to steal confidential information. Many machine learning techniques have been introduced and used because they can analyse a vast data set in a short period and perform a spot-on classification and deduction based on these data. This literature review discusses the approaches and methods researchers have put forward in identifying rogue and phishing sites.

Instead, in their paper [1] "A New Method for Detecting Phishing Websites: URLDetection" (2018), [authors] Shraddha Parekh et al. introduced a new approach for extracting phishing websites through a URL parser; the publication of this method meets the IEEE Xplore standard. By applying a machine learning algorithm to phishing URLs, it attempts to identify any suspicious URL that could be a phishing website.

A threshold is set before passing the sequence to the following layer. Sequence models can also help detect domain names. One helpful idea in this regard, proposed by Sanjukta Mohanty in 2022 [3], would be first to filter the space of one-dimensional features and work with a subset of features deemed relative for training the models. She reported more accurate phishing URL prediction on a one-dimensional base.

Hasane Ahammad Shaik (2022) [5] conducted a study titled 'Robust and effective phishing URL detection using different machine learning methods'. Their research on machine learning-based techniques contributes to the creation of phishing detection systems.

A survey about malicious URL detection proposed with machine learning approaches by D Sahoo (2022) [6] does a great job pinpointing the state of the art and what might be the next research directions on this matter.

In the paper 'Predicting Phishing Sites using Machine Learning: Strategies and Challenges' (2021) [online], Aniket Garje [8]. proposed to detect phishing websites using machine learning. Their technique picks out fake websites by analysing the specifications of websites using machine learning algorithms.

Rakesh Verma et al. (2017) [9] designed a feature extraction and malicious URL detection technique. This technique is said to increase the accuracy of phishing websites. What happens in their approach is extracting features from phishing URLs.

Dipayan Sinha, Ujjwal Mishra [10], Niharika Patnaik and Krishna Kumar K P (2020) proposed an opportune method to detect phishing site URLs and recognise phishing URLs using machine learning. The study on machine learning algorithms in recognising phishing URLs was carried out by author Yadava.

Other machine learning models [12-33] are also referred and helped us to develop our proposed model to give solutions for the identified problem. Taken together, these studies outline the potential of machine learning in mitigating the issue of phishing websites. Driven by a range of diverse methods and techniques, researchers seek to develop resilient and reliable phishing detection mechanisms capable of securing consumers' private data.

III. PROPOSED SYSTEM

The proposed system integrates advanced ML models like Gradient Boosting Classifier, Random Forest, and Decision Tree for precise phishing URL detection. By using 30 attributed feature set and 13 using 13 key features, it distinguishes legitimate URLs from phishing attempts. Real-time protection mechanisms enable swift responses to threats when integrated into web browsers and security applications. Leveraging technologies like NumPy, Pandas, Matplotlib, Scikit-learn, and Flask, the system ensures enhanced cybersecurity defences. Continuous model refinement and feature analysis promise ongoing improvements for heightened accuracy and threat mitigation.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

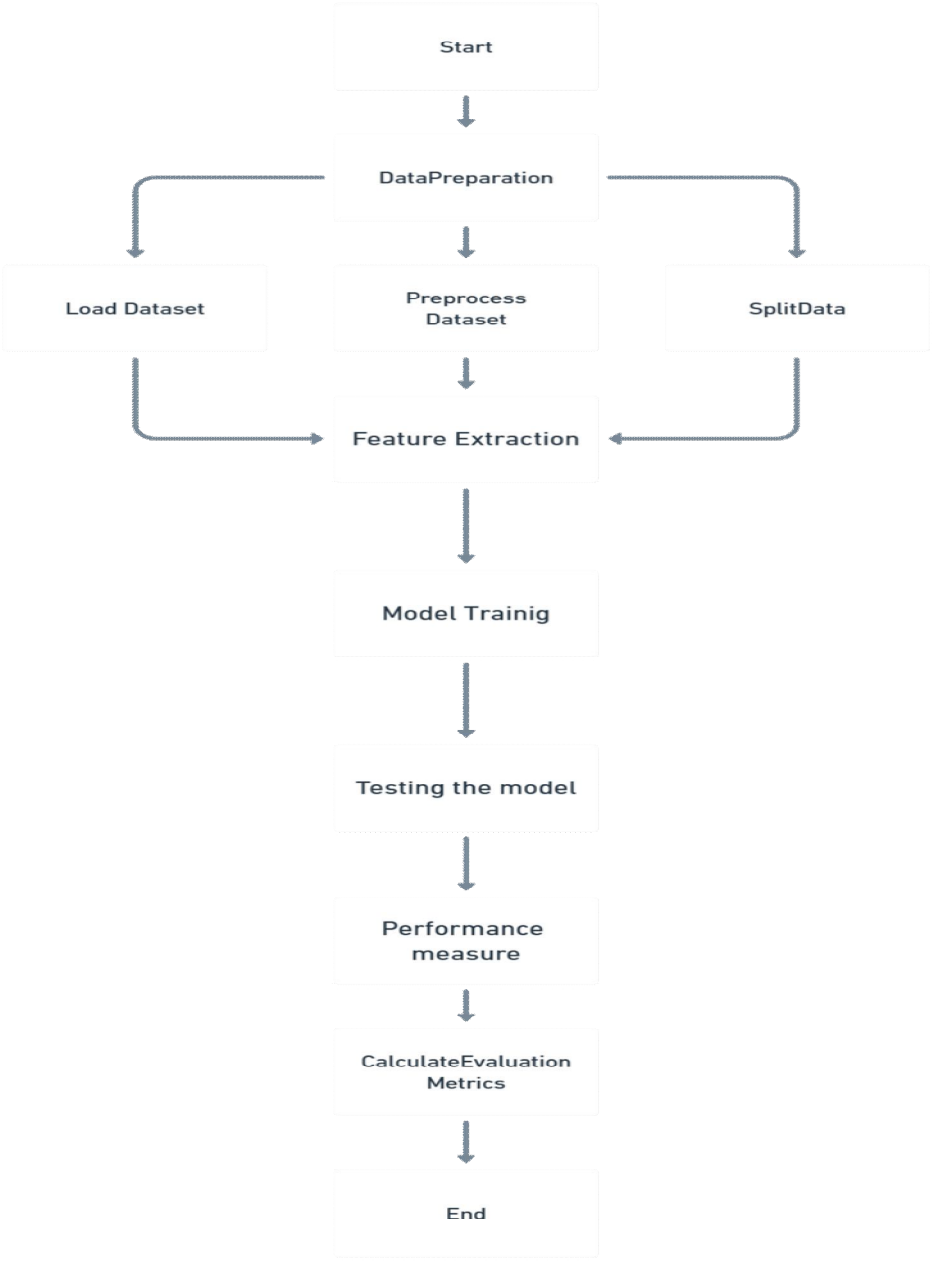


Fig 2. Flow of operations

.

IV. OBJECTIVES

The primary objective of this project is to develop a sophisticated phishing URL detection system utilising Machine Learning (ML) technique. By comparing models (Gradient Boosting Classifier, Random Forest, and Decision Tree) and feature sets, the system aims to classify URLs effectively as legitimate or phishing attempts. Through the extraction of features, the system seeks to enhance cyber security measures.

A. Key objectives include

1) Implementing advanced ML models for accurate phishing URL classification.
2) Extracting crucial features to distinguish between safe and malicious URLs.
3) Comparing 30 attribute feature set and 13 key feature sets.

4) Providing real-time protection against phishing attacks through system integration.
5) Leveraging technologies like NumPy, Pandas, Matplotlib, Scikit-learn, and Flask for efficient system development.
6) Continuously refining ML models and *feature analysis for improved detection capabilities and enhanced cybersecurity defences.*

V. METHODOLOGY

A. Dataset

In our study, we use a dataset borrowed from Kaggle. The dataset contains a collection of website URLs that includes more than 11,000 samples. Each sample is characterised by 30 site parameters and a class label indicating whether it is identified as a phishing site or not, denoted by 1 or -1.

Dataset overview: Total number of samples: 11,054

Properties: 32 (including 30 site parameters)

Target label: Phishing website (1) or Legitimate website (-1)

The dataset serves as a valuable resource for training and evaluating machine learning models for detecting phishing websites. With many samples and a diverse feature set, it offers ample opportunities to explore and implement different algorithms for classification tasks.

B. Data Preprocessing

In our study, the data preprocessing phase played a critical role in ensuring the quality and consistency of the dataset. Initially, we collected a diverse set of features representing various aspects of URL and domain characteristics associated with phishing attempts. This raw dataset underwent thorough cleaning to address issues such as missing values, duplicates, and inconsistencies. We employed techniques such as data imputation, removal of redundant entries, and standardisation of formats to prepare the dataset for analysis. Additionally, we conducted exploratory data analysis to gain insights into the distribution and characteristics of the features, enabling us to make informed decisions during preprocessing. The meticulous preprocessing phase laid the foundation for subsequent model development, ensuring that the dataset was well-prepared for training and evaluation.

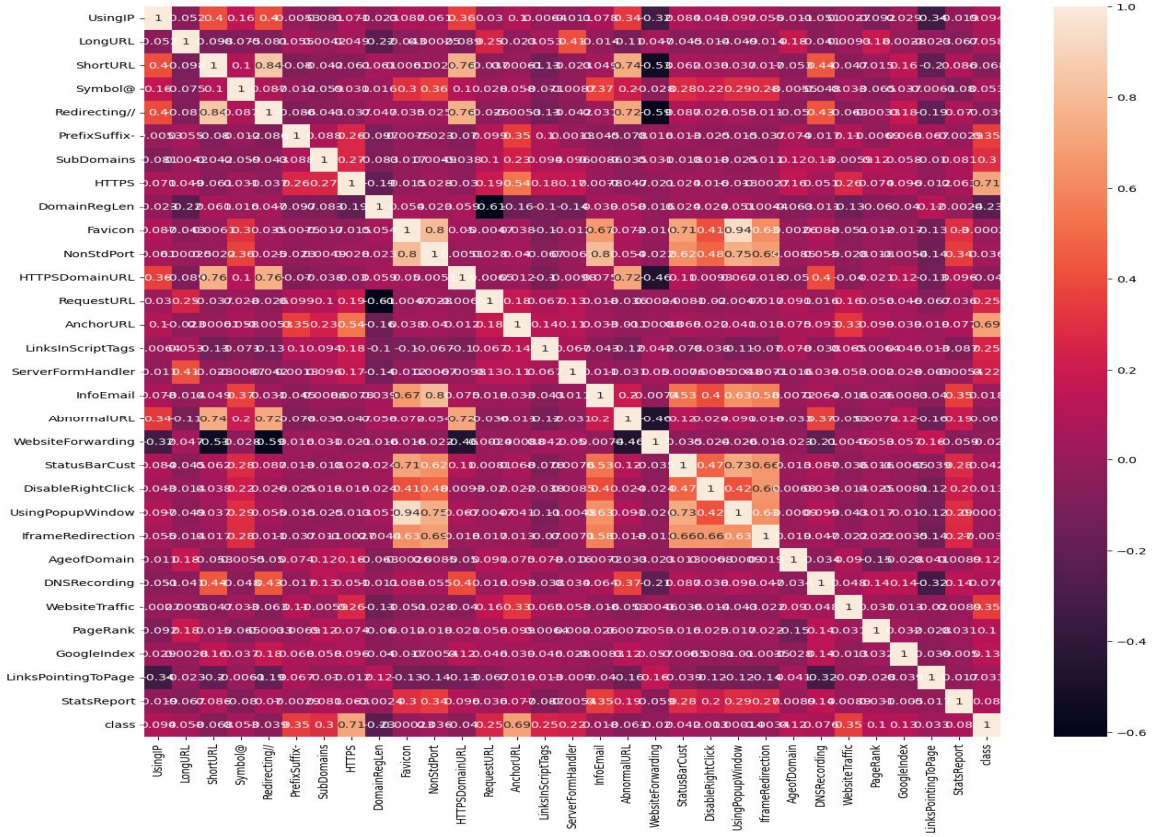


Fig.3. Correlation Heat map

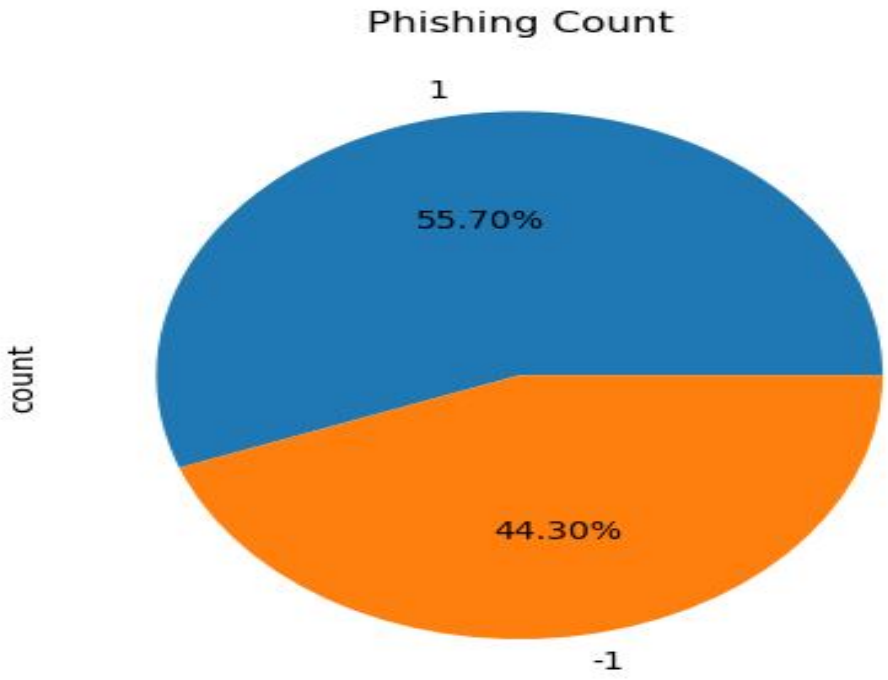


Fig.4: Phishing Count in a pie chart

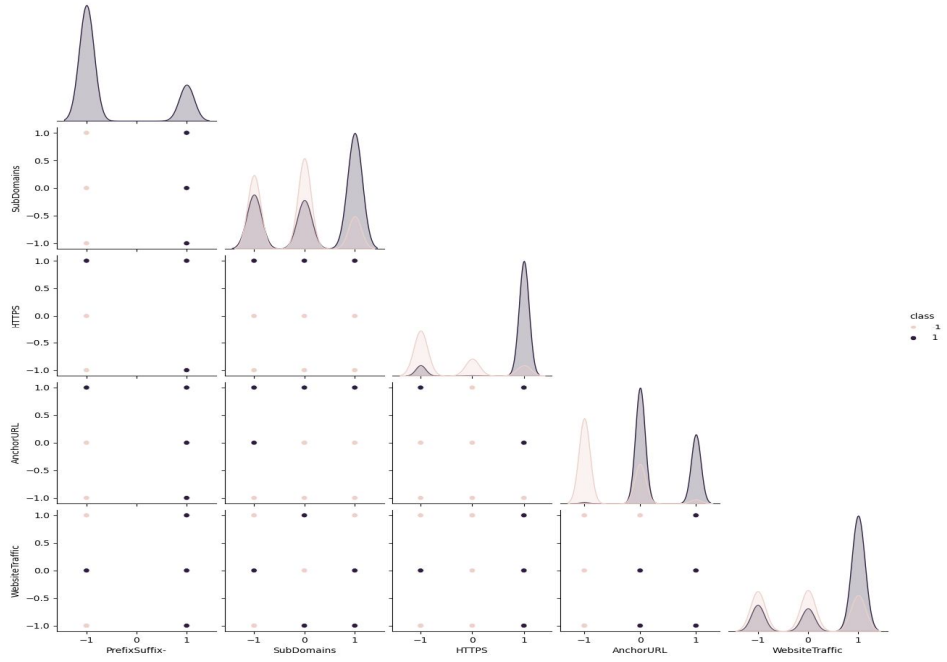


Fig.5: pair plot for particular features

C. ML Model Implementation

1) Importing the necessary libraries: Start by importing the required libraries, including classifiers from the Scikit-learn library, such as "GradientBoostingClassifier," "Random Forest" & "Decision Tree."

2) Split the dataset: Split the dataset into two subsets: one for training the model and one for testing its performance. This ensures independent datasets for training and evaluation using Scikit-learn's train_test_split function.

3) Classifier initialisation and training: Initialise each classifier, including the transition boost classifier, the random Forest classifier, and the decision tree classifier, using the appropriate class constructors. Then, each classifier will be trained using the training data.
4) Make predictions: After training classifiers, use them to make predictions on the test data set. This involves passing the test features to each trained classifier and obtaining the predicted labels.
5) Evaluate the models: Evaluate the performance of each classifier by comparing the predicted labels to the actual labels from the test dataset. Compute evaluation metrics such as accuracy scores and confusion matrices using functions from the Scikit-learn metrics module.
6) Print evaluation results: Print evaluation results for each classifier, including accuracy scores and confusion matrices, to assess their effectiveness in detecting malicious URLs.
7) This Python implementation follows a unified approach to implement machine learning classifiers (Gradient Boosting, Random Forest, and Decision Tree) in a phishing URL detection project, which includes data partitioning, model initialisation, training, prediction, and evaluation steps.

D. Performance Metrics

1) *Accuracy:* Accuracy measures the proportion of correctly classified instances out of the total number of instances.
Formula: Accuracy = Number of correct predictions / Total Number of predictions.
Accuracy ranges from 0 to 1, where 1 means perfect classification and 0 means no correct predictions.
2) *Precision:* Precision measures the proportion of true positive predictions among all positive predictions made by the model.
Formula: Precision = TP / (TP + FP)
Where: FP (False Positives) is the number of negative instances that were incorrectly predicted as positive

3) *Recall:* Recall, also known as Sensitivity or True Positive Rate, measures the proportion of actual positives that are correctly identified by a classification model.
Formula: Recall = TP / (TP + FN)
Where: TP (True Positives) is the number of correctly predicted positive instances.
FN (False Negatives) is the number of actual positive instances that were incorrectly predicted as negative.

4) *F1-Score:* F1-score is the harmonic mean of precision and recall. It provides a balance between precision and recall.
Formula: F1-score = 2 * (Precision * Recall) / (Precision + Recall)
It ranges from 0 to 1, where 1 is the best F1 score.

Briefly, precision measures the overall correctness, precision focuses on the accuracy of positive predictions, recall emphasises the coverage of positive cases, and the F1 score balances precision and recall. These metrics are commonly used to evaluate classification models, including those used in the Phishing URL Detection.

E. Comprehensive Feature set

The comprehensive feature set comprises 30 attributes carefully curated to encompass various dimensions of URL and domain characteristics commonly associated with phishing attempts. Each feature provides valuable insights into different aspects of website behaviour and structure, facilitating the identification of potential phishing threats. The features are:

1) UsingIp
2) LongURL
3) ShortURL
4) Symbol@
5) Redirecting//
6) prefix suffix-
7) SubDomains
8) HTTPS
9) Favicon
10) DomainRegLen.
11) NonStdPort

12) HTTPSDomain
13) AnchorURL
14) LinksInScriptTags
15) DisableRightClick
16) UsingPopupWindow
17) IframeRedirection
18) WebsiteForwarding
19) DNSRecording
20) WebsiteTraffic
21) LinksPointingToPage
22) ServerFormHandler
23) InfoEmail
24) AbnormalURL
25) AgeofDomain
26) StatusBarCust
27) PageRank
28) GoogleIndex
29) Request
30) StatsReport

By considering these comprehensive features, machine learning models can effectively identify patterns and characteristics associated with phishing attempts, thereby enhancing the accuracy of phishing detection mechanisms.

F. Feature Selection

Feature selection is the process of selecting a subset of relevant features (variables or attributes) from a larger set of available features. The goal of feature selection is to improve model performance, reduce overfitting, and reduce computational complexity by focusing on the most informative and discriminating features.

Feature Selection Techniques:

1) SelectKBest:

SelectKBest is a one-dimensional feature selection method that selects the top k features based on statistical tests between each feature and the target variable.

How it works: SelectKBest evaluates each element individually using a statistical test (e.g. chi-square test, ANOVA F-value) and selects the elements with the highest scores.

Usage: SelectKBest is commonly used for classification tasks to help identify the most important features for predicting the target variable.

Example: In Python, SelectKBest can be implemented using the SelectKBest class from the Scikit-learn library.

2) Chi-square test:

The chi-square test is a statistical test used to determine the independence of two categorical variables. When selecting a function, it measures the dependence between each element and the target variable.

How it works: For each element, the chi-square test calculates a chi-square statistic that quantifies the difference between the observed and expected frequencies of each category in the element and the target variable.

Use: The chi-square test is suitable for feature selection when working with categorical data or when the relationship between the features and the target variable is non-linear.

Example: In Python, a chi-square test for feature selection can be performed using the chi2 function from the Scikit-learn library.

In summary, feature selection is a critical step in the machine learning process that helps improve model performance and interpretability by selecting the most relevant features.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

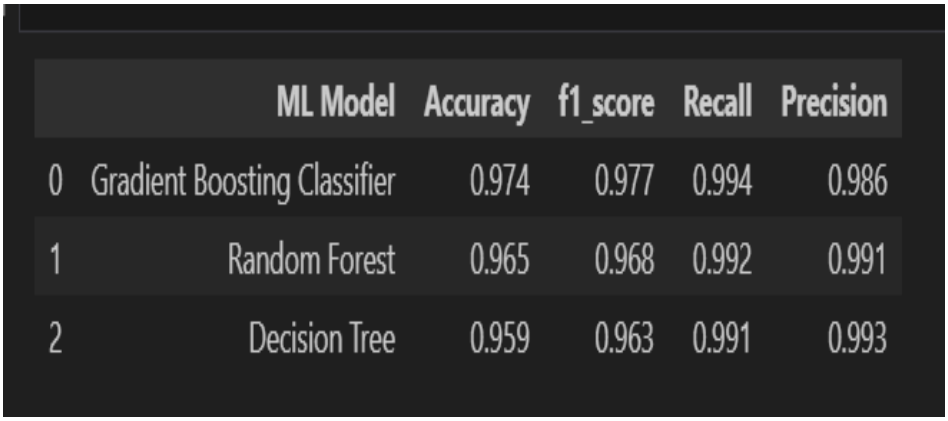
G. Key Feature set

The key feature set comprises 13 attributes carefully selected to capture the most informative indicators associated with phishing attempts. These features are instrumental in identifying patterns indicative of malicious activity and distinguishing between legitimate and suspicious URLs. The key features are:

1) Using IP
2) Long URL
3) Prefix-Suffix
4) Subdomains
5) HTTPS
6) Request URL
7) Anchor URL
8) Links in Script Tags
9) Server Form Handler
10) DNS Recording
11) Website Traffic
12) Google Index
13) Links Pointing to Page

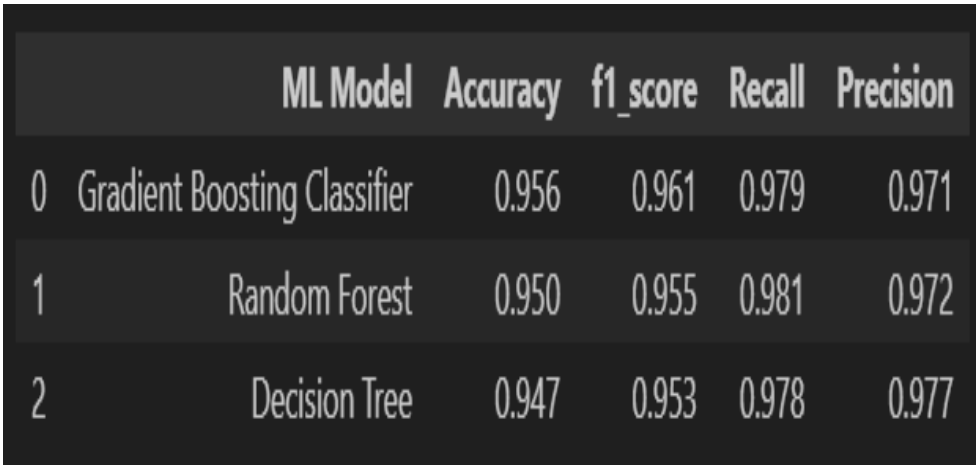
H. Comparing Feature set

Comparison of features entails assessing the performance of models trained with the comprehensive feature set against those trained with the reduced feature set. By identifying key features through feature selection techniques, from the Fig.6 and from the Fig.7 the Fig.6 shows the best Accuracy value, i.e., the Comprehensive Feature Set gains the best accuracy than the key feature set.



	ML Model	Accuracy	f1_score	Recall	Precision
0	Gradient Boosting Classifier	0.974	0.977	0.994	0.986
1	Random Forest	0.965	0.968	0.992	0.991
2	Decision Tree	0.959	0.963	0.991	0.993

Fig.6: Model values on the comprehensive set



	ML Model	Accuracy	f1_score	Recall	Precision
0	Gradient Boosting Classifier	0.956	0.961	0.979	0.971
1	Random Forest	0.950	0.955	0.981	0.972
2	Decision Tree	0.947	0.953	0.978	0.977

Fig.7: Model values on key feature set

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue III Mar 2024- Available at www.ijraset.com

I. Comparing Models

Overall, while Gradient Boosting Classifier often results in high performance across all metrics, Decision Tree and Random Forest can also perform well depending on the dataset and tuning parameters. From the two sets of features, we can see that the Gradient Boosting Classifier gives the best accuracy from Figure Fig.6 and Fig.7.

VI. GRAPHS

1) Decision Tree: plotting the training & testing accuracy for max_depth

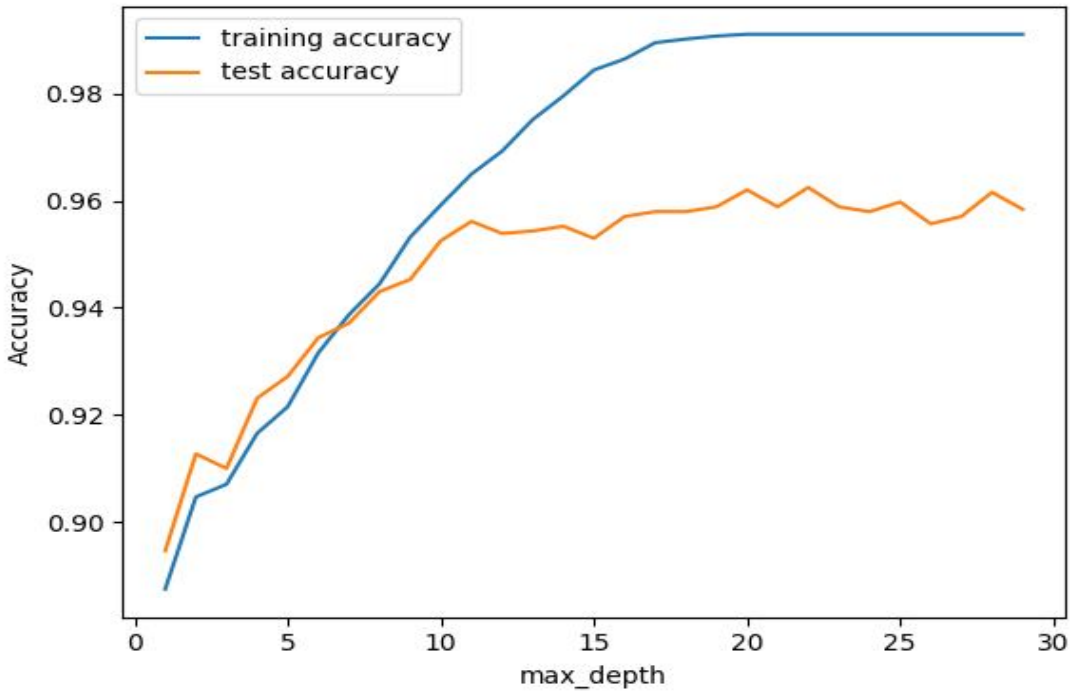


Fig.8: Decision tree

2) Random Forest: plotting the training & testing accuracy for n_estimators

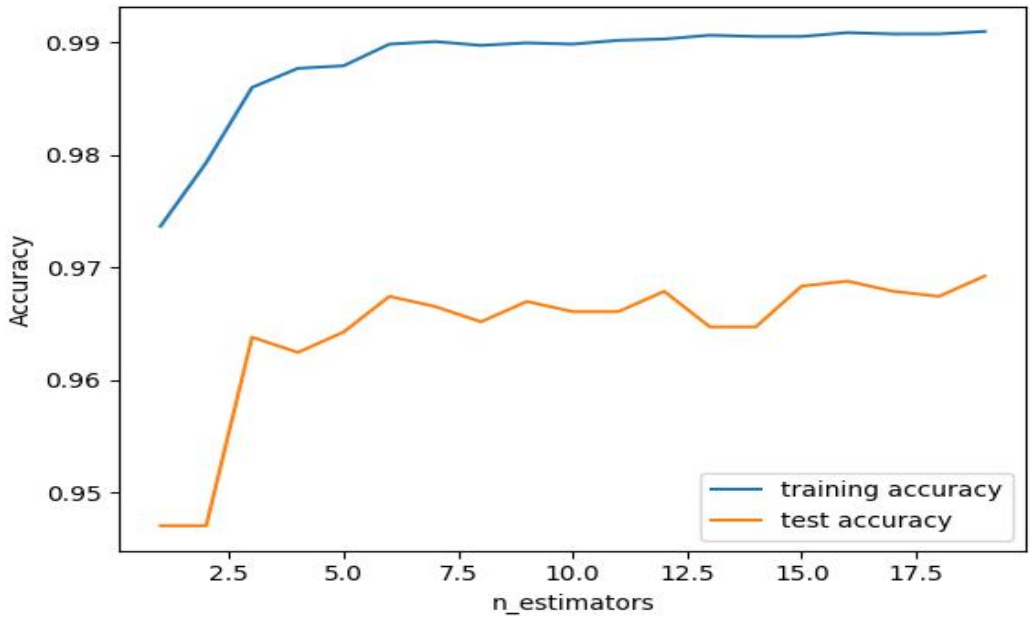


Fig.9: Random Forest

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

3) Gradient Boosting Classifier: plotting the training & testing accuracy for n_estimators

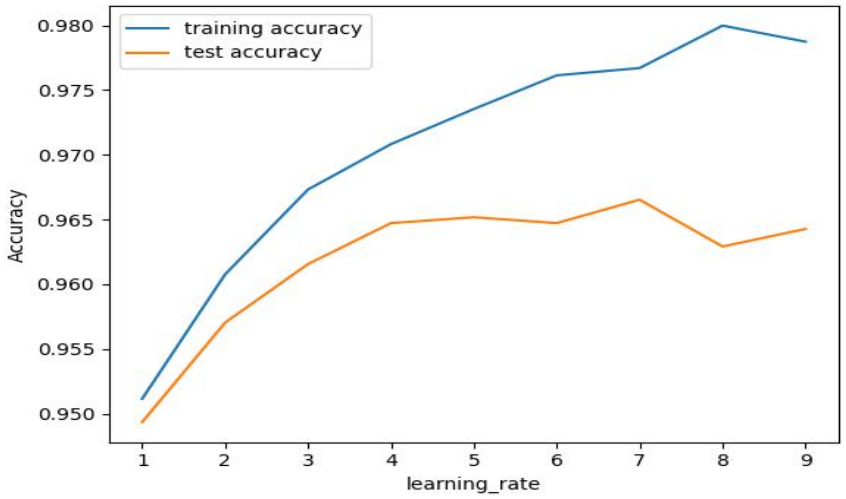


Fig.10: Gradient Boosting Classifier

4) Feature Importances Using Permutation on Full Model:

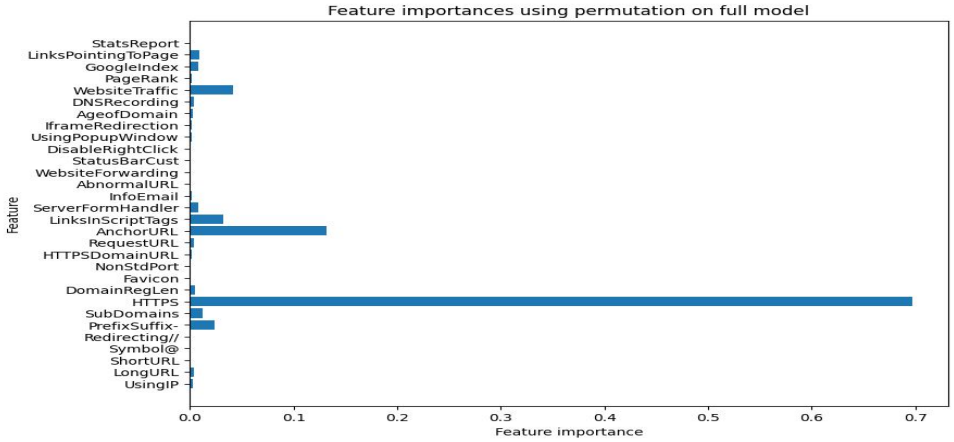


Fig.11: for a comprehensive set of features

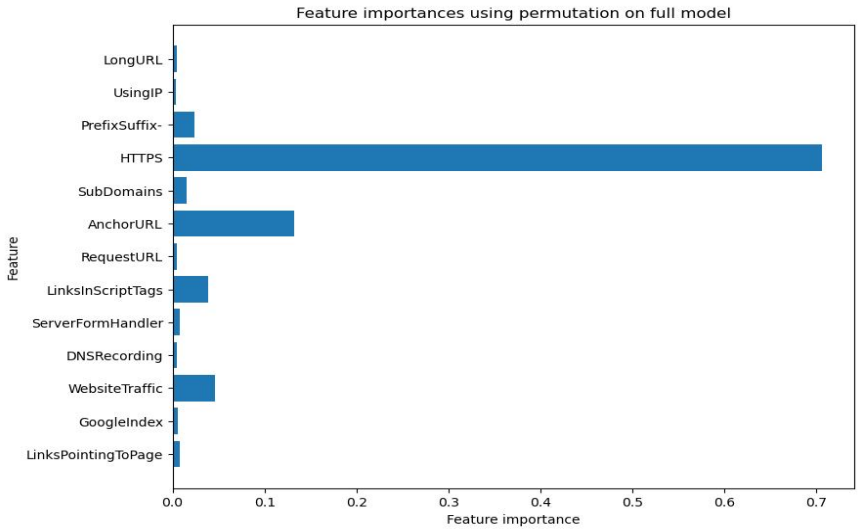


Fig.12: for key features set

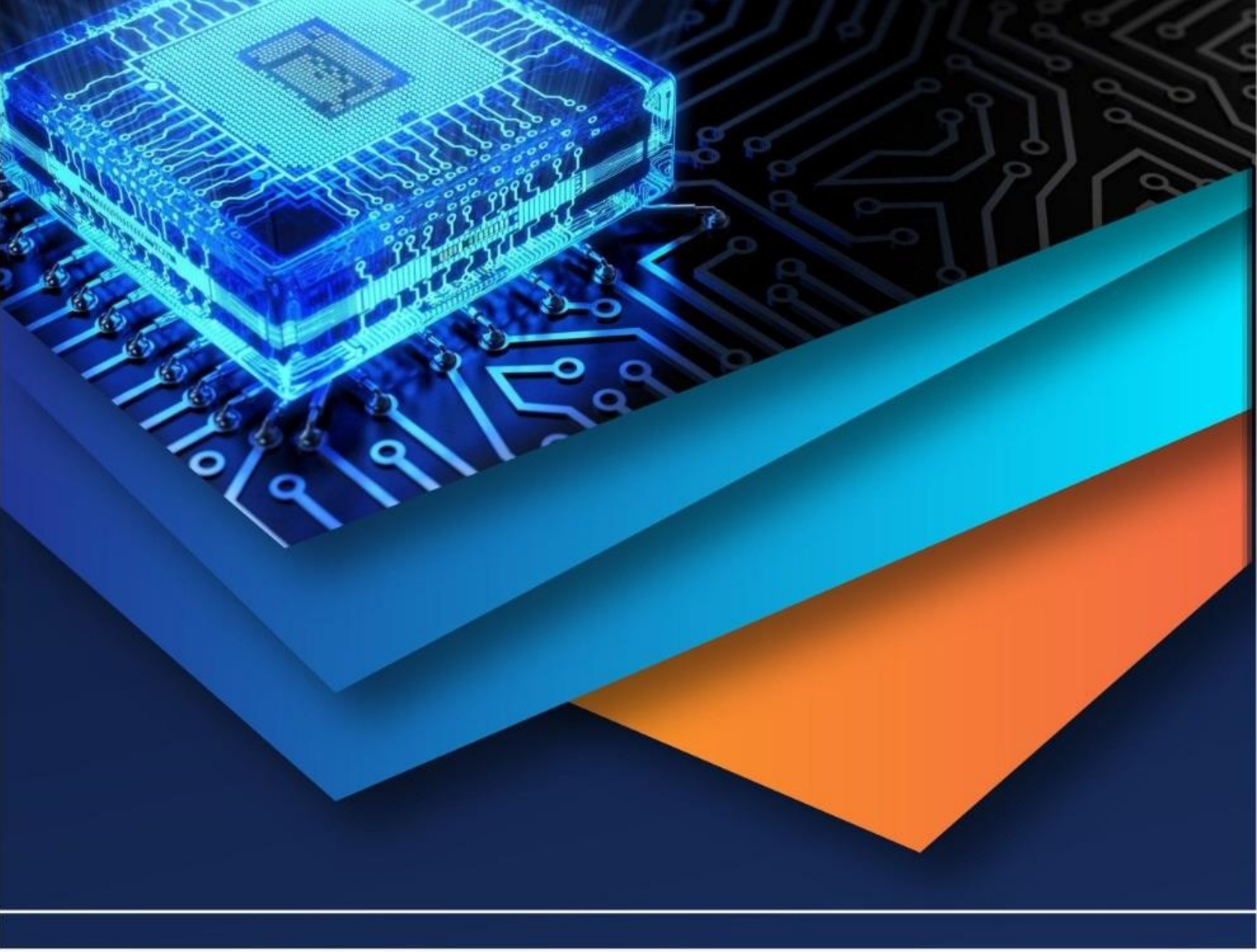
VII. CONCLUSION

In conclusion, our study Successfully developed machine learning models, like Gradient Boosting Classifier, Decision Tree, and Random Forest algorithms, also by considering feature selection techniques, to proficiently detect phishing websites. Through a comprehensive evaluation process, we attained accuracy, precision, recall, and F1-score metrics, showcasing the robustness of our approach. The incorporation of feature selection techniques, such as SelectKBest and Chi-Square, ensured that only the most informative attributes were utilised, enhancing model interpretability and performance. Furthermore, the utilisation of a comprehensive feature set provided invaluable insights into phishing indicators, amplifying the models' detection capabilities. The practical implications of our findings are profound, offering strengthened cybersecurity defences for organisations and individuals alike. Looking ahead, continued exploration into advanced ensemble techniques and the development of real-time detection systems hold promise for further fortifying our defences against the ever-evolving landscape of cyber threats.

REFERENCES

[1] Shraddha Parekh,Dhwanil Parikh, Srushti Kotak, Prof. Smita Sankhe "Title of the article" like "A new method for Detection of Phishing Websites: URLDetection", IEEE Xplore Compliant - Part Number: CFP18BAC-ART; ISBN:978-1-5386-1974-2, Issue 2018
[2] IEEE 2017 - Feature selection for machine learning based detection of phsihng websites
[3] http://ieeexplore.ieee.org/abstract/document/8090317/?reload=true
[4] Sanjukta Mohanty “Predicting Phishing URL Using Filter based Univariate Feature Selection Technique ”, IEEE Conference, Issued on 2022.
[5] V S Lakshmi, M S Vijaya, “Efficient prediction of phishing websites using supervised learning algorithms,” Procedia Engineering, 30, 798-805,2012
[6] Hasane Ahammad Shaik KL University “Phishing URL detection using machine learning methods”, ResearchGate, Issue January 2022.
[7] D. Sahoo, “Malicious URL detection using machine learning: a survey”, 2022.
[8] UpendraShetty D R, AnushaPatil, Mohana,Malicious URL Detection and Classification Analysis using Machine Learning Models,IEEE Xplore Part Number: CFP23CV1-ART; ISBN: 978-1-6654-7451-1, Issue on 2023.
[9] Aniket Garje, Namrata Tanwani, Sammed Kandale, Twinkle Zope, Prof. Sandeep Gore, “Detecting Phishing Websites Using Machine Learning” 2021 IJCRT | Volume 9, Issue 11 November 2021 | ISSN: 2320-2882
[10] Rakesh Verma et al. “What’s in a URL: Fast Feature Extraction and Malicious URL Detection”, Seventh ACM Conference on Data and Application Security and Privacy, pp.55-63,2017.
[11] Dipayan Sinha, Dr. Minal Moharir, Prof. Anitha Sandeep, “Phishing Website URL Detection using Machine Learning,” International Journal of Advanced Science and Technology, vol. 29, no. 3, pp. 2495-2504, 2020.
[12] Sri Hari Nallamala, Dr. Pragnyaban Mishra, KLEF, Dr. Suvarna Vani Koneru, VRSEC, Breast Cancer Detection using Machine Learning Way, International Journal of Recent Technology and Engineering (IJRTE), Vol.8, Issue-2S3, July 2019, ISSN: 2277-3878.
[13] Sri Hari Nallamala, Dr. Pragnyaban Mishra, KLEF, Dr. Suvarna Vani Koneru, VRSEC, Pedagogy and Reduction of K-NN Algorithm for Filtering Samples in the Breast Cancer Treatment, International Journal of Scientific & Technology Research (IJSTR), Vol.8, Issue 11, November 2019, ISSN: 2277-8616.
[14] N B Naidu, Sri Hari Nallamala, Chukka Swarna Lalitha, Syed Seema Anjum, VVIT, Pertaining Formal Methods for Privacy Protection, International Journal of Grid & Distributed Computing,ISSN: 2005-4262, Vol. 13, No. 1, March – 2020.
[15] Kranthi Madala, Sushma Chowdary Polavarapu, VRSEC, Sri Hari Nallamala, Automatic Signal Indication System through Helmet, International Journal of Advanced Science and Technology, Vol. 29, No. 05, April/May – 2020, ISSN: 2005-4238.
[16] Sri Hari Nallamala, Dr. D. Durga Prasad, PSCMRCET, J. Ranga Rajesh, MICT, Dr. Pragnaban Mishra, KLEF, Sushma Chowdary P, VRSEC, A Review on Applications, Early Successes & Challenges of Big Data in Modern Healthcare Management, TEST Engineering and Management Journal,Vol. 83, Issue 3, May – June 2020, ISSN: 0193-4120.
[17] K B Prakash, KLEF, Rama Krishna E, NEC, Nalluri Brahma Naidu, Sri Hari Nallamala, VVIT, Dr. Pragyaban Mishra, P Dharani, KLEF, Accurate Hand Gesture Recognition using CNN and RNN Approaches, International Journal of Advanced Trends in Computer Science and Engineering, Vol. 9, No. 3, May – June 2020, ISSN: 2278-3091.
[18] Sushma Chowdary P, Kranthi Madala, VRSEC, M Sailaja, PVPSIT, Sri Hari Nallamala, Investigation on IoT System Design & Its Components, Jour of Adv Research in Dynamical & Control Systems, Vol. 12, Issue-06, June – 2020, ISSN: 1943-023X.
[19] Sri Hari Nallamala, Bajjuri Usha Rani, LBRCE, Anandarao S, LBRCE, Dr. Durga Prasad D, PSCMRCET, Dr. Pragnyaban Mishra, KLEF, A Brief Analysis of Collaborative and Content Based Filtering Algorithms used in Recommender Systems, IOP Conference Series: Materials Science and Engineering, 981(2), 022008, December 2020, ISSN: 1757-899X.
[20] Manukonda Vinay, Gonugunta Bhanu Sankara Sai Venkatesh, Malempati Venkata Priyanka, Dogiparthi Venkata Sai, Dr. Sri Hari Nallamala, VVIT, Deep Learning Based Face Mask Detection for User Safety from Covid-19, International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE), e-ISSN: 2320-9801, p-ISSN: 2320-9798, Volume 10, Issue 5, May 2022.
[21] P. Radha Vyshnavi, M.V.N. Sai Niharika, M. Summayya, P. Pravallika, Dr. Sri Hari Nallamala, VVIT, Liver Disease Prediction Using Machine Learning, International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET), e-ISSN: 2319-8753, p-ISSN: 2320-6710, Volume 11, Issue 6, June 2022.
[22] Y. Vineela Devi, T.Akshara, S.Mohitha, V.Venkatesh, N.Sri Hari, VVIT, Precision Farming By Analysing Soil Moisture and NPK Using Machine Learning, International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET), e-ISSN: 2319-8753, p-ISSN: 2320-6710, Volume 11, Issue 6, June 2022.
[23] Dr. N. Sri Hari, M. Ramya Sri, Mythri .P, N. Sai Harshitha, M. VenkataNaga Sai Kumar, Detection of Covid-19 using Deep Learning, IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCE, UGC CARE Listed (Group -I) Journal, Volume 11, Iss 12, Dec 2022; P-ISSN: 2319 1775, Online-ISSN: 2320 7876.

[24] Dr. N. Sri Hari, P. Vanaja, M. Ajay Kumar, M.D.V.S. Akash, K. Sivaiah, Multi Disease Detection using Machine Learning, IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCE, UGC CARE Listed (Group -I) Journal, Volume 11, Iss 12, Dec 2022; P-ISSN: 2319 1775, Online-ISSN: 2320 7876.
[25] Dr.N.Sri Hari, Shaik Nelofor, Siramdasu Leela Vardhan, Sura Rana Prathap Reddy, Sakhamuri Devendra, CycleGAN Age Regressor, International Journal for Innovative Engineering and Management Research, ISSN: 2456-5083, Volume 12, ISSUE 04, Pages: 45-51, April 2023.
[26] Sudheer Mangalampalli, Ganesh Reddy Karri, Amit Gupta, Tulika Chakrabarti, Sri Hari Nallamala, Prasun Chakrabarti, Bhuvan Unhelkar, Martin Margala, Fault-Tolerant Trust-Based Task Scheduling Algorithm Using Harris Hawks Optimization in Cloud Computing, Sensors 2023, 23(18), 8009; https://doi.org/10.3390/s23188009.
[27] K. Sudharson, …, Sri Hari Nallamala, et al., Hybrid Quantum Computing and Decision Tree based Data Mining for Improved Data Security, 7th International Conference On Computing, Communication, Control And Automation (ICCUBEA), August 18-19, 2023, 979-8-3503-0426-8/23/$31.00 ©2023 IEEE.
[28] G Sanjay Gandhi, K Vikas, V Ratnam, K Suresh Babu,IET Communications 14 (16), 2840-2848 ,Grid clustering and fuzzy reinforcement-learning based energy-efficient data aggregation scheme for distributed WSN
[29] KV Prasad, GS Gandhi, S Balaji , Inexpensive colour image segmentation by using mean shift algorithm and clustering,, International Journal of Graphics and Image Processing 4 (4), 260-266
[30] PSK venkatraomaddumala, sanjay gandhi gundabatini, P AnushaClassification of Cancer cells detection using Machine Learning Concepts,International Journal of Advanced Science and technology 29 (3), 9177-9190
[31] Sanjay Gandhi Gundabatini, Suresh Babu Kolluru, C. H. Vijayananda Ratnam & N. Nalini Krupa ,. DAAM: WSN Data Aggregation Using Enhanced AI and ML Approaches Conference paper First Online: 27 June 2023,LNEE,volume 976
[32] Gundabatini, S.G. Rayachoti, E., Vedantham, R Recurrent Residual Puzzle based Encoder Decoder Network (R2-PED) model for retinal vessel segmentation. Multimed Tools Appl (2023). https://doi.org/10.1007/s11042-023-16765-0
[33] Gundabatini, S.G.Rayachoti, E., Vedantham, R EU-net: An automated CNN based ebola U-net model for efficient medical image segmentation. Multimed Tools Appl (2024). https://doi.org/10.1007/s11042-024-18482-8
[34] MANUEL SÁNCHEZ-PANIAGUA , EDUARDO FIDALGO FERNÁNDEZ , ENRIQUE ALEGRE, WESAM AL-NABKI , AND VÍCTOR GONZÁLEZ-CASTRO, Phishing URL Detection: A Real-Case ScenarioThrough Login URLs, IEEE ACCESS date of publication April 18, 2022
[35] DataSet: https://www.kaggle.com/eswarchandt/phishing-website-detector
[36] MoitrayeeChatterjee, AkbarSiami Namin, Detecting Phishing Websites through DeepReinforcement Learning, 2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)