Classification of Features for detecting Phishing Web Sites based on Machine Learning Techniques

Sandeep Kumar Satapathy, Shruti Mishra, Pradeep Kumar Mallick, Lavanya Badiginchala, Ravali Reddy Gudur, Siri Chandana Guttha

Abstract: ***Phishing is one of the most common and dangerous attacks among cybercrimes. The aim of this attack is to hack the user information by accessing the credentials that is used by individuals and any of the organizations. Phishing websites contents and web-based information contains various hints. The victim's confidential data is expected by the phishing sites by deriving them to surf a phishing website that resembles to legitimate website, which is one of the criminal attacks prevailing in the internet. Phishing websites is similar to cyber threat that is targeting to get all the credential-based information such as information accessed from the credit cards and social security numbers. Till now there is no specific solution that can detect phishing attacks and also truly unpredictable which includes numerous components and also criteria that are not stable. The purpose of this project is to perform Extreme Learning Machine (ELM) based classification for 30 features including Phishing Websites Data in UC Irvine Machine Learning Repository database. There are different types of features based on web pages. Hence, to prevent phishing attacks we must use a specific web page feature. We proposed a model based on machine learning techniques like Naïve Bayes to detect phishing web pages. For results assessment, ELM was compared with other machine learning methods such as Naïve Bayes (NB), ANN and detected to have the highest accuracy of 89.3%.***

Index Terms*: *Extreme Learning Machine, Features Classification, Information Security, Phishing

I. INTRODUCTION

Phishing is a Web-based attack that seduces end users to visit fake websites and give away personal information such as user id and password. Phishing web pages are formed by fraudulent people to copy a web page from an original one. These phishing web pages are very similar to the original ones. Technical tricks and social engineering are extensively joined together for beginning a phishing attack. An important view of online security is to protect users from phishing attacks and fake website. Intelligent methods can be used to develop fake web pages. For this reason, internet users whether have enough experience in information security or not might be cheated. Phishing attacks can be launched via sending an e-mail that seems to be sent from a trusted public or private organization to users by attackers. Attackers get the users to update or verification their information by clicking a link within the e-mail. Other methods such as file sharing, blogs, and forums can be used by attackers for phishing. There are many ways to fight phishing including legal solutions, education, and technical solution. Nowadays, information and communication tools are used in a manner that is very dense with information. For this purpose, various solution methods for various problem types have been developed.

Machine Learning (ML) methods, can also be used in application development for information security. Optimization, classification, prediction and decision support system and great benefits can be provided to the person who is responsible for information security. There are attacks for different purposes to the Information and Communication tools that create computer networks. These attacks can be detected and the necessary precautions should be taken. For the study of artificial intelligence seems to gain speed as computer technology evolves. Artificial intelligence methods and studies on information security are increasing day by day. Intelligent systems provide great benefits in deciding to information security professionals. ML methods can be used with classification purposes in various fields. Classification can be considered as a process to determine whether a data belong to one of the classes in the dataset organized according to certain rules.

Revised Manuscript Received on May 22, 2019.

Sandeep Kumar Satapathy, Department of Computer Science & Engineering, KL Deemed to be University, Vijayawada, Andhra Pradesh, India
Shruti Mishra, Department of Computer Science & Engineering, Vellore Institute of Technology, Amaravati, Andhra Pradesh, India
Pradeep Kumar Mallick, School of Computer Engineering, Kalinga Institute of Industrial Technology (KIIT) Deemed-to-be University, Bhubaneswar, Odisha, India
Lavanya Badiginchala, Undergraduate Students, Department of Information Technology, Vignana Bharathi Institute of Technology, Hyderabad, India
Ravali Reddy Gudur, Undergraduate Students, Department of Information Technology, Vignana Bharathi Institute of Technology, Hyderabad, India
Siri Chandana Guttha, Undergraduate Students, Department of Information Technology, Vignana Bharathi Institute of Technology, Hyderabad, India
Email: sandeepkumar04@gmail.com

Classification which used in many fields and has an important place has a separate place for information security.

Neural nets models have been used in many areas such as data mining, medical applications, chemical industry, energy production, electrical and electronics industry, communications, nonlinear system modelling, pattern matching.

In this study, an intelligent model for detecting phishing web pages based on Extreme Learning Machine is presented. We have suggested some new rules to have efficient features. The average classification accuracy as a result of the tests 95.05% evaluated. The paper is organized as follows, at first a brief of introduction for the study and related works about different phishing detection techniques are represented. Secondly the phishing threat, Extreme Learning Machines and details about the dataset that is used in intelligent model are summarized. Thirdly rules of used features and k-fold cross validation test briefly explained. Fourthly application of intelligent model is given in details. At last conclusions are given.

It is intended to explain and prove the important feature, and prediction of the websites in the source number. In addition, some new features were proposed. Experimentally, it was appointed the new rules for some well-known properties. Updates were made to some other features.30 rules created for the attributes of the prepared data set examined.

Using the IP address: Feature 1: As an alternative, an IP address in the URL domain name can be used. Sometimes an IP address can be converted into radix 16 codes.

Rule: If IP address exists in the domain →phishing, otherwise → legitimate

URLlength: Feature 2: The average URL length has been calculated. If the number of URL characters is equal to 54 or greater than 54 then URL has been classified as phishing.

Rule: If the URL length < 54 → legitimate, URL length ≥ 54 and ≤ 75 → suspicious, otherwise → phishing

Using TinyURL: Feature 3:URL length can be shortened and even a web page can be opened in this way. Short URL domain name, which depends on behalf of the Long URL domain, can be performed with HTTP Redirection.

Rule: IfTinyURL is containing in it → phishing, otherwise → legitimate

Using "@" symbol: Feature 4:It's been aforesaid that succeeding a part of"@" symbol in URL is ignored by the browser. It has been said that the next part of "@" symbol in URL is often the real address.

Rule: If URL is containing @ symbol → phishing, otherwise → legitimate

Using "//" symbol: Feature 5:The user may be directed to another web site using "//"in URL.If URL starts with "HTTP" then "//" symbol must be in the 6th position. If URL starts with "HTTPS" then "//" symbol must be in the 7th position.

Rule: The position of last occurrence of "//" in URL > 7 → phishing, otherwise→ legitimate.

Using "-" symbol: Feature 6:The dash symbol is rarely used in the legitimate URL.In this way users think that they are using a legitimate web page.

Rule: If (underscore)"-" symbol exists in domain name → phishing, otherwise → legitimate

Sub and multi sub domain: Feature 7:"www." and country code in the URL are ignored. The remaining points are counted in the URL.

If Dots in Domain part is equal to1 then it is Legitimate and if the Dots in Domain part are 2 the it is Suspicious otherwise phishing website.

Rule: number of dots in domain = 1 → legitimate, number of dots in domain = 2 → suspicious, otherwise → phishing

Using HTTPS: Feature 8: The authors have been suggested checking the certificate including HTTPS used, trusted certificate issuer, and the certificate age.

Rule: Using HTTPS, trusted security certificate providers, age of certificate ≥ 1 year → legitimate

Using HTTPS, untrusted security certificate providers → suspicious, otherwise → phishing

Domain registration length: Feature 9:It has been found that the fake domains which is longest have been used for one year only in the dataset.

Rule: domains expires on ≤ 1 year → phishing, otherwise→ legitimate

Favicon: Feature 10:If a web page that contains the favicon is loaded from a domain different from the domain shown in the address bar, then the web page has been classified as "phishing".

Rule: favicon loaded from external domain → phishing, otherwise → legitimate

Standard port status: Feature 11: It has been investigated open or closed status of the service on a server with this feature. The port number, service name, description, and preferred status are shown in the Table5regarding some of the ports that are used in general.

Using HTTPS token: Feature 12: HTTPS token can be added to a part of domain of URL by attackers.

Rule: In domain part of URL using HTTPS token → phishing, otherwise → legitimate

Request URL: Feature 13:Web page address and most of the objects which are embedded in web pages may share the same domain in a legitimate web page.

Rule: % of request URL<22%→ legitimate, % of request URL >=22% and <61% → suspicious, otherwise → phishing

URL_of_anchor: Feature 14: Anchor has been identified as a member indicated by tag. tags and the web site may have different domain

names. The anchor element may not be a connection to any web page.

Rule: % of URL_of_anchor<31%→ legitimate, % of URL_of_anchor>=31% and <=67%→ suspicious, otherwise → phishing

Links in <meta>, <script>ve<link>, Feature 15: These tags are expected to be connected to the same domain on a web page. Tag is used to retrieve metadata about the HTML (Hyper Text Mark-up Language) document recommendation. <Script> tag is used to create client-side script. tag is used to get other web resources

Rule: % of links in <meta>, <script> and <link>tags<17%→ legitimate, % of links in <meta>, <script> and <link> tags >=17%and 81%→ suspicious, otherwise → phishing

Server Form Handler: Feature 16: SFH (Server Form Handler) that contain an empty string or about: blank classified as "phishing". If the domain name in SFH is different from the domain name of the webpage, then classified as "suspicious".

Rule: SFH is "about: blank" or empty → phishing, SFH refers to a different domain → suspicious, otherwise → legitimate

Submitting information to e-mail: Feature 17: A web form is used to send a user's personal information to a server. "mail ()" function can be used by using a server-side language and "mailto" can be used by using a client-side language.

Rule: using "mail ()" or "mailto:" → phishing, otherwise → legitimate

Abnormal URL: Feature 18: This feature could be extracted from the WHOIS database. Identity is typically part of its URL for a legitimate website.

Rule: Host name is not in URL → phishing, otherwise → legitimate

Website forwarding: Feature 19:It has been found that legitimate websites are redirecting mostly once, and phishing websites are redirecting at least 4 timesin the dataset.

Rule: number of redirect page ≤ 1 → legitimate, number of redirect page ≥ 2 and < 4 → suspicious, otherwise → phishing

Status bar customization: Feature 20: A fake URL can be displayed to the users in the status bar by the attackers. JavaScript can be used for this purpose. Especially "onMouseOver" event was focused on.

Rule: onMouseover changes status bar → phishing, otherwise → legitimate

Disabling right click: Feature 21: JavaScript can be used for this purpose. The source code of a web page could not be displayed and recorded by the user in this way. "event. Button==2" event has been investigated in a source code of webpage.

Rule: right click disabled → phishing, otherwise → legitimate

Using pop-up window: Feature 22: Request to send the users' personal information in a pop-up window on a legitimate website is not regarded as a normal situation. This feature can be used in some legitimate websites for specific purposes.

Rule: popup window contains text field → phishing, otherwise → legitimate

Iframe redirection: Feature 23: It has been said that to show an extra webpage the iframe tag is used.

Rule: using iframe → phishing, otherwise → legitimate

Age of domain: Feature 24: This feature could be extracted from the WHOIS database. It is observed that an age of legitimate domain is at least 6 months.

Rule: age of domain ≥ 6 months → legitimate, otherwise → phishing

DNS record: Feature 25:An identity of phishing website is not recognized or no records are found for the host name in the WHOIS database. If the DNS (Domain Name System) record does not exist or has not been found, then website is classified as "phishing". Otherwise it is classified as "legitimate".

Rule: no DNS record for domain → phishing, otherwise → legitimate

Website traffic: Feature 26: This feature is measured interest in a website. Because of phishing websites live for a short period of time they may not be recognized by the Alexa database. It was found that the legitimate websites are among the top in the ranking of 100. 000.If the domain has no traffic or it is not recognized by the Alexa database, then it has been classified as "phishing". Otherwise it has been classified as "suspicious". The values of Alexa Traffic Ranks are shown for http://www.ucla.edu/ website. The Traffic Ranking values were measured for Global and The United States in 2026 and 662 respectively

Rule: website rank < 100.000 → legitimate, website rank > 100.000 → suspicious, otherwise → phishing

PageRank: Feature 27: It has been said that PageRank is a value from 0 to 1. It has been found that 5% of phishing webpages may reach a PageRank value up to "0.2". The values between 0 and 1 in the PageRank algorithm, the values between 1 and 10 in the Google Toolbar PageRank tool are used

Rule: PageRank < 0,2 → phishing, otherwise → legitimate

Google Index: Feature 28: A site is displayed on search results when it is indexed by Google. Because of phishing webpages that can be accessed for a short period generally, many phishing webpages may not be found in the Google Index.

Rule: webpage indexed by Google → legitimate, otherwise → phishing

Number of links pointing to page: Feature 29: This feature has been defined

about legitimate level even if some links are on the same domain. It has been observed that legitimate websites have at least 2 external links pointing to them in the dataset.

Rule: number of links pointing to webpage = 0 → legitimate, number of links pointing to webpage> 0 and ≤ 2 → suspicious, otherwise → legitimate

Statistical reports: Feature 30: Many statistical reports on phishing websites have been defined for period of times by Phish Tank and Stop adware. Two types of top ten of Phish Tank have been used in the study. These types are "top 10 domains" and "top 10 IPs". "Top 50 IPs" of Stop adware have been used.

Rule: host in top 10 phishing IPs or domains → phishing, otherwise → legitimate

II. RELATED WORK

Based on structural properties in phishing email detection the proposed approach briefly explains to find phishing through proper and appropriate identification and usage of structural propertied of email. This project is done by ANN, Naïve Bayes and classification technique to classify phishing emails. The classification method is not large enough and uses one approach to detect suspicious emails which is low in efficiency and scalability. This technique is based on email structural properties and has to extend more content properties in order to reduce the error results during analysis. By using some techniques and algorithm like intelligent phishing website detection and prevention system that uses link guard algorithm, a system using link guard that works for hyperlinks has been proposed.

Certain tests like comparisons of DNS of actual and visual links were performed by the algorithms and also checks dotted decimal of IP address, checks coded links and pattern matching. The system contains some of the drawbacks that it produces false positive results if any legitimate and trust worthy site has IP address instead of domain name and if the user does not visit the legitimate (original) site then it considers some suspicious sites as normal. Hence the assessment results in false negative conclusions.

The Statistics of suspicious URL's have been analyzed by many researchers. The garera uses to classify phishing website URL's the work by garera uses logistic regression over hand selected features. It includes features like flag keywords in URL, Google page rank-based features web page quality guidelines. A direct comparison with our approach is difficult without access to same URL's and features.

A comparative analysis of phishing website and non-phishing website URL did not construct a classifier by MC Grath and Gupta. But they compared from DMOZ open directory project drawn from non-phishing URL's to phishing URL's in the phish Tank. The analysis features include IP address who is in the records containing time and date information about geographic, registration and also lexical features of URL such as length, width, distribution of characters and predefined popular names.

III. METHODOLOGY

A. Artificial Neural Network

ANN is used in order to understand the impact of increasing or decreasing the dataset vertically and horizontally in dynamic time. It leads to the difficult task which to specify the network architecture that is necessary in terms of the number of hidden layers and the number of neurons in each hidden layer associated with building a neural network model and it contains a set of parameters (learning rate, momentum, epoch size) should be specified in advance in order to build a good model. Unfortunately, it is difficult to identify the appropriate network structure for a particular application, and that could be reached by trial and error.

Procedure1: Assigning the random weights is necessary to start the algorithm using the inputs and that maps to the hidden nodes find the activation function for each of the hidden nodes using activation rate of the hidden nodes and links to the output find the activation rate of the output nodes. Error rate at the output node needed to be found using the weights and error found at output node record the error at the hidden nodes. If the actual output is not similar to the target output then backtrack and modify the weights till, we reach target output.

Accuracy of ANN: 87.901%

B. Naïve Bayes

It is a classification technique supported by Bayes' Theorem with an assumption of independence among predictors. In easy terms, a Naive Bayes classifier assumes that the presence of a specific feature in an exceedingly class is unrelated to the presence of another feature. For example, a fruit is also thought of an apple if 's red, round, and about three inches in diameter. Even if these features rely upon each other or upon the existence of the opposite features, of all those properties severally contribute to the probability or likelihood that this fruit is an apple and hence it is known as 'Naive'.

Procedure2: There are three types of naïve bayes classifiers. Naive bayes classifier while dealing with real time data which has continuous distribution considers that the data which is generated is to be big through the gaussian process with normal distribution. First, we need to train the data applying some naïve bayes classifier builder to get appropriate naïve bayes classifier. The outcome model will have the high

performance with gaussian naïve bayes classifier that contains high training speed with capabilities to predict the capability of the feature that belongs to zk classifier.To compute the ith observation would be by computing the following probability.

Probability of (zk| x(i)).
Then applying naïve bayes rule it can be written as prob(zk|x(i)) =prob(zk)prob(x(i)|zk/prob(x(i)).

Accuracy of Naive Bayes: 61.365%

C. Extreme Learning Machines

Extreme learning machines are unit feed forward neural networks for classification, regression, clustering, sparse approximation, compression and have learning with one layer or multiple layers of hidden nodes, where the parameters of hidden nodes (not simply the weights connecting inputs to hidden nodes) needn't be tuned. These hidden nodes is discriminately allotted and never updated (i.e. they are random projection however with nonlinear transforms), and are often inherited from their ancestors while not being modified.

Procedure3

ELM algorithm: Randomly generate hidden node parameters and randomly assign hidden nodes. (wi,bi), where i=1,2,3,…L; calculate output matrix of hidden layer. Then calculate output weight matrix H.

$$\hat{\beta} = H^{\dagger}T$$

IV. EXPERIMENTAL RESULTS

As observed extreme leaning machine acquires the highest measurable values when compared to others.

Table I. The below table shows the precision, recall, f1score and accuracy for three different Machine learning techniques.

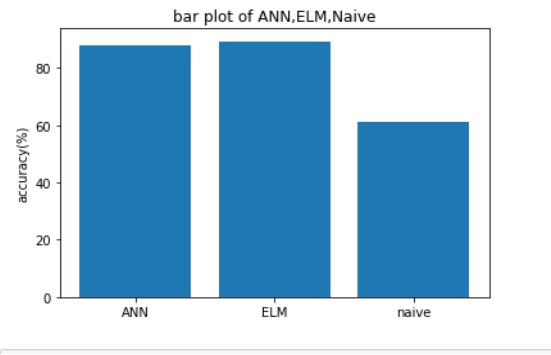

Fig. 1. Graph to show accuracy for three algorithms

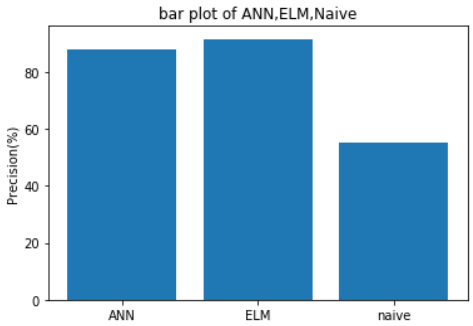

Fig. 2. Graph to show precision for three algorithms

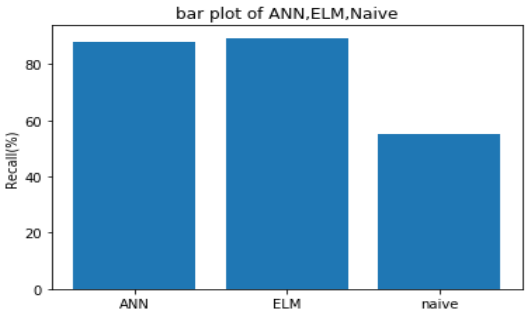

Fig. 3. Graph to show recall for three algorithms

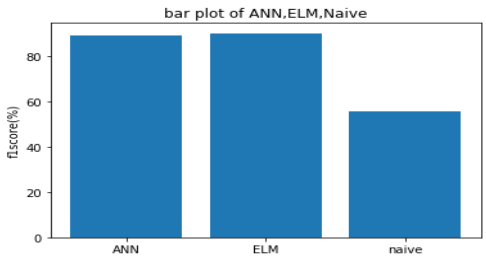

Fig. 4. Graph to show f1score for three algorithms

ML Techniques	Precision (%)	Recall (%)	F1score (%)	Accuracy (%)
Naive Bayes	55.6	55.5	55.6	61.3
ANN	89.3	89.3	89.3	87.9
ELM	91.5	89.3	90.4	89.3

As observed in all the above graphs that containing precision, recall, f1score and accuracy. Where ELM is having highest values when

compared to ANN. Naïve Bayes is acquiring very low values.

Accuracy and f1 score for ELM Activation Functions

Table II. The below table shows the accuracy and f1score for each of the activation functions of ELM

Measures	Accuracy	F1score
Sigmoid	89.1	89.9
Tanh	90.2	89.2
Sine	56.4	61.6
Tribas	66.4	70.3
Multiquadric	83.0	84.5
Inv_tribas	66.2	72.5
Inv_multiquadric	72.8	82.3
Hardlim	84.2	86.9
Softlim	89.3	88.3
Gaussian	71.2	75.6

Predictions

First, we need to give the URL which is provided by the users. Let us check the URL Structure based on important characteristics like Identity of domain, Security and Encryption criteria in the final phishing detection rate.

For Example:
https://www.exampleurl.com is the given URL so we need to check whether the given website is phishing or non-phishing.

Steps:
Input: enter the URL.
Process: Check the URL based on given feature conditions.
Output: After checking the conditions if it returns -1 then the provided URL is Non-Phishing else if it provides other than -1 it is phishing website.

V. CONCLUSION

Systems varying from data entry to information processing applications can be made through websites. The entered information can be processed; the processed information can be obtained as output. Nowadays, web sites are used in many fields such as scientific, technical, business, education, economy, etc. Because of this intensive use, it can be also used as a tool by hackers for malicious purposes. One of the malicious purposes emerges as a phishing attack. A website or a webpage can be imitated by phishing attacks and using various methods. Some information such as users' credit card information, identity information can be obtained with these fake websites or web pages.

The purpose of the application is to make a classification for the determination of one of the types of attacks that cyber threats called phishing. Extreme Learning Machine is used for this purpose. In this study, we used a data set taken from UCI website. In this dataset, input attributes are determined in 30, and the output attribute is determined in 1. Input attributes can take 3 different values which are 1, 0, and -1. Output attribute can take 2 different values which are 1, and -1. *k*-fold cross validation test has been implemented where k=10, for measuring the performance of generated system in this study. As a result of the study, the average classification accuracy was measured as 95.05%, and its highest accuracy was to be measured as 95.93%. When the dataset is examined, it has been observed that the rule created for feature 13 where are classified in the form of legitimate, suspicious, and phishing. When the dataset was examined by us, it was observed that 13th attribute values were consisted of 1 and -1. It was detected by us that the 13th attribute has 6560 legitimate and 4495 phishing samples which are 11055 samples totally.

ACKNOWLEDGMENT

This work has been completed with the support of DST – FIST sponsored lab in Vignana Bharathi Institute of Technology, Hyderabad, India.

REFERENCES

1. P. Ying and D. Xuhua, "Anomaly based web phishing page detection, "in Proceedings - Annual ComputerSecurity Applications Conference, ACSAC, 2006, pp. 381–390.
2. M. Moghimi and A. Y. Varjani, "New rule-based phishing detection method," Expert Syst. Appl., vol. 53, pp. 231–242, 2016.
3. DATASET: Lichman, M. (2013). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science
4. G.-B. Huang et al., "Extreme learning machine: Theory and applications," Neurocomputing, vol. 70, no. 1–3, pp. 489–501, 2006.
5. C. S. Guang-bin Huang, Qin-yu Zhu, "Extreme learning machine: A new learning scheme of feed forward neural networks," Neurocomputing, vol. 70, pp. 489–501, 2006
6. T. S. Guzella and W. M. Caminhas, "A review of machine learning approaches to Spam filtering," Expert Systems with Applications, vol. 36, no. 7. pp. 10206–10222, 2009.
7. W. D. Yu, S. Nargundkar, and N. Tiruthani, "A phishing vulnerability analysis of web based systems," IEEE Symp. Comput. Commun. (ISCC2008), pp. 326–331, 2008.
8. P. Ying and D. Xuhua, "Anomaly based web phishing page detection,"in Proceedings - Annual Computer Security Applications Conference,ACSAC, 2006, pp. 381–390.
9. M. Moghimi and A. Y. Varjani, "New rule-based phishing detectionmethod," Expert Syst. Appl., vol. 53, pp. 231–242, 2016.
10. S. K. Satapathy, S. Dehuri, A. K. Jagadev, "An Empirical Analysis of Different Machine Learning Techniques for Classification of EEG Signal to Detect Epileptic Seizure", International Journal of Applied Engineering Research

ISSN 0973-4562 Volume 11, Number 1 (2016) pp 120-129, Research India Publications.
11. T. S. Guzella and W. M. Caminhas, "A review of machine learning approaches to Spam filtering," Expert Systems with Applications, vol.36, no. 7. pp. 10206–10222, 2009.
12. G.-B. Huang et al., "Extreme learning machine: Theory and applications," Neurocomputing, vol. 70, no. 1–3, pp. 489–501, 2006.
13. C. S. Guang-bin Huang, Qin-yu Zhu, "Extreme learning machine: Anew learning scheme of feed forward neural networks, Neurocomputing, vol. 70, pp. 489–501, 2006.
14. S. K. Satapathy, S. Dehuri, A. K. Jagadev, "An Empirical Analysis of Training Algorithms of Neural Networks: A Case Study of EEG Signal Classification Using Java Framework", Springer India 2015, L.C. Jain et al. (eds.), Intelligent Computing, Communication and Devices, Advances in Intelligent Systems and Computing 309, DOI 10.1007/978-81-322-2009-1_18
15. Ö. Faruk Ertugrul and Y. Kaya, "A detailed analysis on extreme learning machine and novel approaches based on ELM," Am. J. Comput. Sci Eng., vol. 1, no. 5, pp. 43–50, 2014.
16. Ö. F. Ertugrul, "Forecasting electricity load by a novel recurrent extreme learning machines approach," Int. J. Electr. Power Energy Syst., vol. 78,pp. 429–435, 2016.

AUTHORS PROFILE

Sandeep Kumar Satapathy is currently working as an Associate Professor in the Department of Computer Science and Engineering, KL Deemed to be University, Vijayawada, Andhra Pradesh, India. He was the former Head of the Department of Information Technology, Vignana Bharathi Institute of Technology, Hyderabad, India. He has completed his doctoral degree in the area of machine learning and data mining. He has been actively involved in the research in the area of Machine Learning, Deep Learning, Computational Intelligence, Image Processing, Data Mining and has more than 45 papers to his credit. He has authored 3 books and has been guest editor for several edited books.

Shruti Mishra is currently working as an Assistant Professor in the School of Computer Science and Engineering, Vellore Institute of Technology, Amaravati, Andhra Pradesh, India. She has completed her doctoral degree in the area of Data Mining and Bioinformatics. She has been actively participated in all the research activities. She has three books to her credit, several book chapters and research papers to her credit. She has also been the guest editor to several edited books.

Pradeep Kumar Mallick, currently working as an Associate Professor, School of Computer Engineering, Kalinga Institute of Industrial Technology (KIIT) Deemed-to-be University, Bhubaneswar, Odisha, India. He was earlier working as the Professor and Head of the Department, Computer Science and Engineering, Vignana Bharathi Institute of Technology, Hyderabad, India. He has completed his doctoral degree in Data Mining and is current pursuing in post-doctoral fellowship in the area of Image Processing. He has many books and has more than 45 research papers to his credit. He has also guest edited several edited books.

Lavanya Badiginchala, currently an undergraduate B.Tech (IT) student in Vignana Bharathi Institute of Technology, Hyderabad, India.

Ravali Reddy Gudur, currently an undergraduate B.Tech (IT) student in Vignana Bharathi Institute of Technology, Hyderabad, India.

Siri Chandana Guttha, currently an undergraduate B.Tech (IT) student in Vignana Bharathi Institute of Technology, Hyderabad, India.