ANALYSIS OF E-COMMERCE USABILITY AND SECURITY FACTOR BASED ON MULTIMODAL BUSINESS DIMENSION IN PHARMACEUTICAL SUPPLY CHAIN AND CBIR BASED TEXTILE RECOMMENDATION SYSTEM

Biresh Kumar*[1], Sharmistha Roy[2], Anurag Sinha[3], Namit Garg[4], Hassan Raza Mahmood[5], Muneeb Hashmi[6]

[1-2]Department of Computing and Information Technology, Usha Martin University, Ranchi, Jharkhand, India.
[3]Department of Computer Science, IGNOU, New Delhi, India.
[4]Department of Maths and Computing, Indian Institute of Technology, Kanpur, Utter Pradesh, India.
[5]Department of Computer Science, FAST NUCES CFD Campus Faisalabad, Punjab, Pakistan.
[6]Department of Computer Science and Engineering, National University of Science and Technology, Islamabad, Pakistan.
[1]bireshmtech@gmail.com, [2]sharmistha@umu.ac.in, [3]anuragsinha257@gmail.com, [4]namitg16@gmail.com, [5]hassanraza999@hotmail.com,[6]muneebhashmi10@gmail.com.
Corresponding Author: Biresh Kumar, Department of Computing and Information Technology, Usha Martin University, Ranchi, Jharkhand, India.
Email: bireshmtech@gmail.com

DOI: 10.47750/pnr.2022.13.S07.336

Abstract

The usability testing is a complex procedure with many components for those wishing to improve their websites' usability to the fullest. There is no rocket science involved; all you need to know is some basic assumptions. The most obvious one is that if you want your website to be fantastic, you must test. The key usability issues that are pertinent to e-commerce websites are linked to the trouble spots during the purchasing process. Consequently, the following inquiries are covered by the exam objectives: Is it simple for customers to locate the items they need? Do they have access to enough information to make a well-informed purchasing decision? What challenges arise? A carefully defined aim would have a quantifiable outcome to be attained (such as work completion rate, duration, accuracy, etc.), a particular. In this hypothetical article, a theory that hasn't been confirmed is presented, and usability testing and security testing are briefly contrasted. Despite the absurdity of making comparisons between these occupations, the result is surprising. Testing may not be as different between usability and security as it first seems. More research on usability testing may yield fresh security testing ideas.

Keywords: E-commerce, usability testing, security testing, machine learning, supply chain, CBIR.

Introduction

It could be challenging to define usability. A system's usability, which is a distinct aspect of software quality, relates to how well it communicates with its users. Usability is typically defined in terms of the sub-aspects that it includes, i.e., by its position in a conceptual tree or as a superset of more specific aspects. Usability is a general concept that encompasses learnability, efficiency (of the interface), memorability, a low error rate, and subjective user enjoyment [2]. Security might be difficult to describe. Security, which may be divided into a variety of smaller variables including dependability, availability, confidentiality, integrity, and the absence of particular flaws, is another element that affects the quality of software. During the testing procedure, many other problems might occur. When too much information is mixed, a user usually ends up injured and bewildered. The visual clutter on the page has to be reduced or the relevant links within the overall hierarchy need to be made more evident in this case. If a user cannot find the category they are seeking for, the information is probably classified differently than they anticipate. Finally, it's conceivable that the user just does not get the idea of the website. Become calm. It's as conceivable for the person taking the test after you to understand every detail in a standard test [1].

Websites for electronic commerce (E-commerce) have increased dramatically over time. However, only 29% of users turn their internet search into a purchase owing to security and usability problems. The purpose of this study is to give a summary of the benefits and drawbacks of the current evaluation methodologies for an e-commerce website's usability and security. This study examined the E-commerce assessment methods used from 2000 to 2018. The study examined 11 models and identified the usability and security components assessed by each model. Each model's advantages and disadvantages were emphasised. This study discovered that there isn't a single complete model that can measure all the usability factors at once [2-5].

E-commerce testing aids in evaluating, confirming, validating, and testing e-commerce websites and apps. It generally examines the app or website's performance, usability, security, compatibility, and functionality. Additionally, e-commerce mobile app testing aids in determining whether the app complies with all requirements for safe transactions in terms of standards, protocols, and laws.

E-commerce testing's primary goal is to make sure:

- The e-commerce website's pages are all routinely examined, and all transactions are fully verified and safe.
- There are no compromises made regarding any significant elements, including the security of client data, responsiveness on mobile devices, loading speed, user experience, etc.
- ensuring that the website or programmed is sufficiently protected against outsiders and online hackers by testing, validating, and confirming.
- that the specific e-commerce programme is error-free, secure, and user-friendly

Software security is a component of computing that is becoming more and more crucial, yet far too many development projects still treat it as an afterthought. Numerous methods have been put out for security requirements engineering, however we find that many of them still have usability issues. In this proposal, I outline the security requirements engineering work I have done. SURE, or Secure and Usable Requirements Engineering, is a novel method that helps non-security specialists express security requirements such that testing artefacts may be created as a result. The SURE approach is also implemented by the system ASSURE, Automated Support for Secure and Usable Requirements Engineering. The major contribution of this paper is to propose ml based model to test security and usability of several business usable models. The paper is arranged as Section 1 of the paper comprises the introduction and significance of the proposed method. The second section of the paper consists of the algorithmic background of the literature review. Section 3 of the paper consists of background and the introduction of the algorithm used in the paper. Sections 4 and 5 contain the method and results, including the conclusion [6-10].

Background

Supply chain model of pharmaceutical industry

One of the most important elements in the automation and optimization of supply-chain operations is radio frequency identification (RFID) technology. Particularly in the pharmaceutical supply chain, RFID tags have the potential to significantly increase inventory management efficiency. As a result, improving medication monitoring in international pharmaceutical production facilities will pave the way for extensive Internet of Things usage

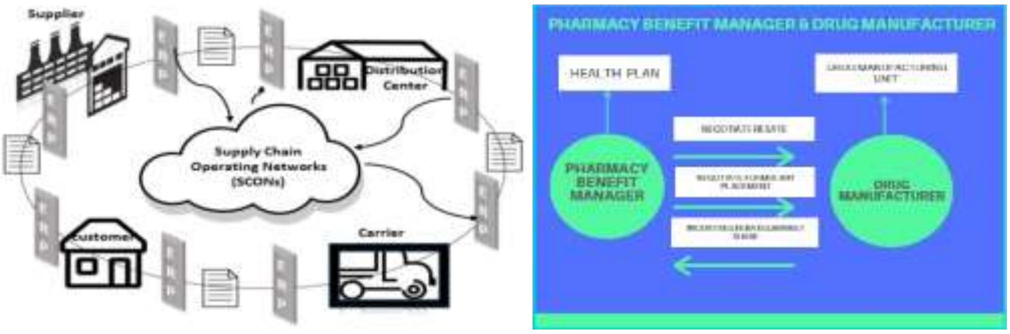
(IoT). As the prevalence of unlicensed pharmacies and fake pharmaceuticals has increased, so have related elements in production processes. To track and regulate the distribution and quality of pharmaceuticals, these sensors continually give a sizable quantity of contextual data, including places and times. Business logic may be successfully applied by using this contextual knowledge. putting together, simulating, analysing, and sharing [11-12]

Table 1 - Top 15 synergistic drug pair predicted by the co-gene/GS score in the DREAM dataset

Drug pair	**Co gene/GS score**	**Predicted rank**	**Gold standard rank**
Camptothecin & Mitomycin C	0.084	1	5
Camptothecin & Doxorubicin	0.036	2	16
H-7 & Mitomycin C	0.032	3	2
Methotrexate & Mitomycin C	0.030	4	28
Doxorubicin & Mitomycin C	0.027	5	4
Cycloheximide & H-7	0.027	6	9
Camptothecin & Etoposide	0.025	7	15
H-7 & Trichostatin A	0.019	8	19
H-7 & Rapamycin	0.017	9	27
Camptothecin & H-7	0.015	10	10
Etoposide & Mitomycin C	0.015	11	3
H-7 & Vincristine	0.012	12	43
H-7 & Monastrol	0.012	13	14
Cycloheximide & Methotrexate	0.011	14	64
Doxorubicin & H-7	0.011	15	1
Drug Combination	**Co gene/GS score**	**Rank**	
Prestwick-682 & MG-262	0.667	1	
Cefazolin & Nocodazole	0.667	2	
Anisomycin & Prednisolone	0.625	3	
Lisinopril & Suramin sodium	0.600	4	
Iopanoic acid & Butacaine	0.563	5	

Our study focuses on the creation of a decision support system for evaluating organisational performance using the expenses associated with hospital logistics. The article's first section, which deals with hospital logistics, provides examples of prior studies on how to assess performance in the commercial and healthcare sectors. The second describes our system for gauging financial performance. An application of our methods to a real-world scenario is focused on the third. We conclude with some career opportunities.

Fig 1; supply chain based pharmaceutical e-commerce link



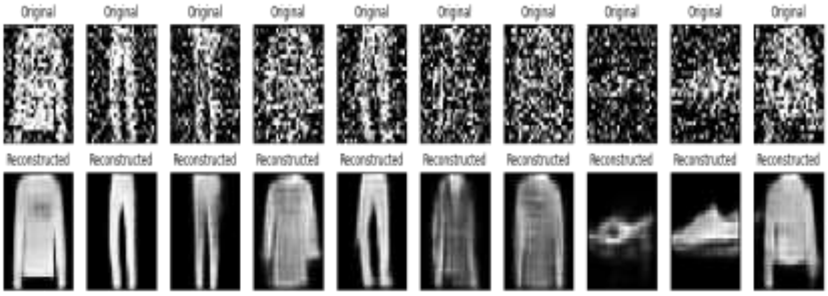
B2B E-Commerce

Business-To-Business is a situation where one business makes a commercial transaction with another. In other words, it is a transaction between one business and another, such as a wholesaler and retailer.Business-to-business refers to a business that is conducted between companies, rather than between a company and individual consumers. It typically occurs when a business is sourcing materials for its production process for output. B2B transactions tend to happen in the supply chain, where one company will purchase its raw materials from another and will use them in the manufacturing process.B2B e-commerce, short for business-to-business electronic commerce, is the sales of goods or services between businesses via an online sales portal. In general, it is used to improve the efficiency and effectiveness of a company's sales efforts. Because orders are processed digitally, buying efficiency is improved for wholesalers, manufacturers, distributors, and other types of B2B sellers. A 2018 survey found that 48% of companies now conduct 50–74% of all corporate purchases online. Additionally, 23% of companies do 75% or more of their purchasing online Emerging eCommerce technologies are also reducing the barrier to entry for traditionally B2C businesses to add a B2B component (B2C2B) and, vice versa, for traditionally B2B companies to sell direct-to-consumer (B2B2C) [13-15].

CBIR based textile recommendation model

Making business decisions is one of the essential elements of any type of business for its productivity and market penetration. Business choices and analytics are crucial in industries like e-commerce and the fashion one. This research seeks to present a solution for business decision-making based on deep learning and ECG classification linked with a language model. In this research, we suggest the Denoising Auto-Encoder based textile image retrieval for figuring out the preferences of the online client based on several personalities. This model utilises the Amigos personality data set in addition to the MNIST data collection for training. The crucial step in content-based picture retrieval is feature extraction. The features lay forth our high-level representation of a picture. Using deep learning algorithms is a different strategy. With the help of this implementation, we can show how to train a convolutional neural network (CNN) for classification tasks and derive a "neural code" for the photos in our database. These neural codes encodedly represent an image's characteristics. The issue with this strategy is that in order to train the neural network, we first require labelled data. An excellent solution to this problem is provided by the Denoising Autoencoder. An autoencoder is a feed-forward neural network that compresses the input into a "code" and then reconstructs the code back into the original input. Since our dataset consists of images, it makes sense to use convolution layers for encoding and decoding. The encoder is stacked with Conv2D and MaxPooling2D layers (the max-pooling layer down samples the spatial dimensions of the images by taking the maximum value over an input window for each channel of the input). The decoder is stacked with Conv2D and UpSampling2D layers (which up-sample the input by doubling the dimensions of the input image). The model converges to a loss value of 0.2489 and is reasonably successful in encoding the images and reconstructing those encoded images as shown below in Figure 4. In order to test the trained auto encoder in a practical scenario, a noise factor is added to the images to corrupt them slightly and see if the encoder can recover the original images. The noised images are as below in Figure 2. The Auto encoder is fitted onto the noisy data, and prediction is made accordingly. As shown below, the model is reasonably successful in finding the right kind of clothing even with much noise as shown in Figure 2 [16-17].

Fig. 2 Predicted Samples



Methodology

Proposed statistical environment

A statistical technique called multiple regression, or MLR, integrates a number of explanatory variables to predict the result of a response variable. Multiple linear regression is used to try to show the linear connection between the explanatory (independent) and response (dependent) variables. Because it takes into account several explanatory variables, multiple regression is only an extension of ordinary least-squares (OLS) regression.

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_p X_{ip} + e$$

Elastic net regression, also known as ELNET regression improves statistical model regularisation by combining LASSO and ridge regression penalties. A sparse model is produced during regularisation by the L1 component of the penalty. The L2 quadratic component of the penalty, on the other hand, promotes the grouping effect and removes the number constraint on the accessible variables while increasing the stability of the L1 part along the path to regularisation. As a result, certain features' impact is lessened while some of them are maintained.

$$\hat{\beta}(Enet) = \left(1 + \frac{\lambda_2}{n}\right)\left\{arg\ min_{\beta} \|y - X\beta\|^2 + \lambda_1 \|\beta\| + \lambda_2 \|\beta\|^2\right\}$$

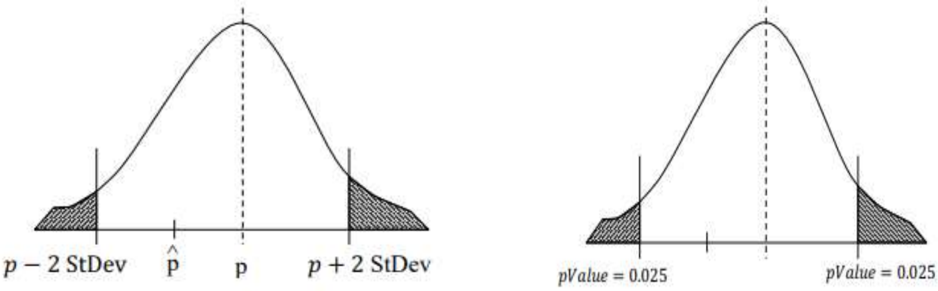
To measure the strength of a linear link between two variables, x and y, correlation coefficients are utilised. A connection is positive if the linear correlation coefficient is higher than zero. There is a negative association when the number is smaller than zero. The variables x and y are said to be unrelated if their value is 0.

$$\text{Correlation}= \rho=\frac{cov(X,Y)}{\sigma_X \sigma_Y}$$

The dependent variable for this study is nut yield, whereas the independent factors are plant height, stem girth, female flowers in inflorescence, husk thickness, copra content, minimum temperature, rainfall, and relative humidity. To identify the elements that are all very much to blame for multicollinearity amongst predictor variables, correlation between nut yield and every other factor is conducted. Given that you have chosen a confidence level of 95%, you anticipate that a portion of the sample (p) will fall between the ranges of (population proportion (p) - 2 StDev p p + 2 StDev 15 Basics of Data Science 2'(Standard Deviation) and (population proportion (p) + 2'(Standard Deviation), . In this time period, there is a 95% chance of occurrence (Please refer to Figure 6). Since there is a 95% chance of success. Also keep in mind that you are guessing p's value and would be computing since you are unsure about it. The value of p will fall between (- 2'(Standard Deviation)) and (+ 2'(Standard Deviation)).

After that, you do the exploratory analysis and create several charts to understand the data's nature. After that, a statistical test for significance is run, such as a chi-square, independent sample t-test, ANOVA, or nonparametric test, depending on the sample size, type, and other factors and data properties. The null hypothesis is assumed to be true in these tests. The p-value, which serves as evidence against the null hypothesis, is a parameter value that a test may provide based on the sample and the probability [18].

Figure 3 ; distribution of data for usability and security testing



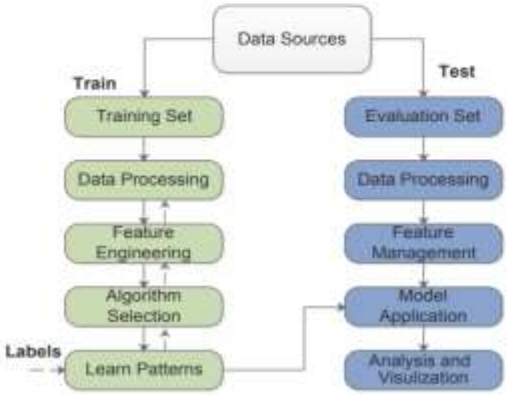
Proposed computing environment

In this section, we give a high-level overview of the multi-agent system paradigm we suggest. A proposed approach for attaining high levels of interaction is to use a range of agent designs (simple, learner and model-based agent). A multi-agent system can achieve its distinguishing attribute of cooperating and working in sequence by breaking down the overall job (classification) into smaller, more manageable jobs and distributing them to agents (MAS). In this study, a highly interactive MAS was created. The five agents that make up this MAS each have their own unique architecture and work together to meet the system's goals and score highly. Three separate classification algorithms were employed to accomplish the assignment, producing a system with a wide range of classifications with varied degrees of accuracy that may be utilised to enhance the decision-input quality. maker's The data from three separate agents, each of which has run its own algorithm, may be used to significantly increase the DM agent's prediction capacity. The first learner will utilise the Random Forest method to build a training model for itself. With the aid of this model, we are able to forecast and obtain a unique result, X (Algorithm (3.4)). The second learner creates its own training model using naive Bayes. With the aid of this model, we may perform the prediction process and generate an original output Y [19-22].

Input: array of features
Output: ***X***

receive dataset from preprocessing agent.
reading dataset.
set a last attribute as a class.
convert class attribute from numeric to nominal.
split the dataset to training dataset & testing dataset.
specifying the number of instances for training.
creating size of instances for training (0-train size).
creating size of instances for testing (train size -end)
set the classifier (rf) for Random Forest.
set the parameters for Random Forest for training.
Training.
Testing.
evaluation
Saving T. model in DB.
Receiving array of features from DM agent.
Predicting
Send the result of predicting to DM aent

Figure 4 ; proposed method



This algorithm has a number of useful characteristics that could be used in discrete time. It is also a dynamic system that generates complex sequences that behave randomly easily and simply, its signal is not random but it is deterministic, it has a high sensitivity of initial conditions so any change in initial conditions creates another sequence, and finally it has random behaviour in the particular space [24]. The 3D Logistic Formulation logistic map that the formula as stated as follows model. Using this model, we can perform the prediction process and get a unique output Y . xi+1 = λxi (1-) + βyi 2 + xi + azi 3 (7) yi+1 = λxi (1-) + βzi 2 + xi + axi 3 (8) zi+1 = λxi (1-x) + βxi 2 + xi + ayi 3 (9)

Security testing

The DHKE is suggested to be updated in this study. This modification was based on a novel optimization problem that was developed utilising the optimised discrete logarithm functions (ODLFs), which are given in the following description(The Optimized Discrete Logarithm Function) (The Optimized Discrete Logarithm Function). Let g be a generator element of the prime field Fp and let h be a nonzero element of Fp. The ODLF function finds the greatest or minimum exponent for the bivariate function a1x1+a2x2 such that (g) max or min(a1x1+a2x2) h. (mod p). Here is a description of the simplex-based Optimized Diffie-Hellman Key Exchange (ODHKE). using as public domain inputs the huge prime p and the generator element g of a prime [23-25].

	x_1	x_2	S_1	S_2	...	S_n	R.H. S
S_1	c_{11}	c_{12}	1	0	...	0	b_1
S_2	c_{21}	c_{22}	0	1	...	0	b_2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
S_n	c_{n1}	c_{n2}	0	0	...	1	b_n
a	$-a_1$	$-a_2$	0	0	...	0	0

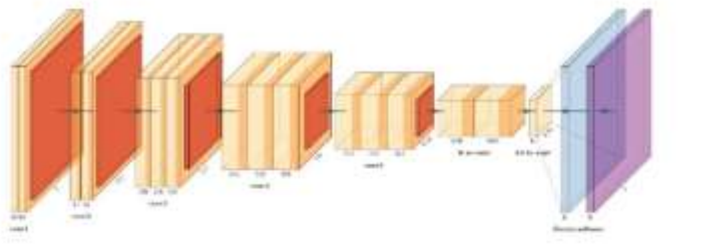
Table 2. A Initial values of objective function and constraints for Bob

	x_1	x_2	S_1	S_2	...	S_n	R.H. S
S_1	c_{11}	c_{12}	1	0	...	0	b_1
S_2	c_{21}	c_{22}	0	1	...	0	b_2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
S_n	c_{n1}	c_{n2}	0	0	...	1	b_n

In this study, a convolutional neural network and long short-term memory network-based intrusion detection and classification model is suggested. The design primarily consists of four parts: data gathering, data processing, classification model using convolutional neural networks and long short-term memory networks, and judgment-based decision-making. In Figure 5, the preprocessed data are first flattened into a 1-dimensional array, then the data are mapped from low-dimensional to high-dimensional through two layers of 3*3 convolution operations, and the height*width is 6*6 channels, feature maps of 64 and 128, respectively, in which the meaning of the data represented by different dimensions is also different; finally, after the maximum pooling operation, the size of the pooling kernel is 2*2, and the height*width is 6*6 channels.

The convolutional layer, which comes before the connection layer in the CNN component, is in charge of extracting the data's features. The conventional approach involves connecting the completely connected layer after gathering the features, followed by activation classification, but the fully connected layer also includes training and testing computations. The number of parameters affects speed, and if there are too many, it is simple to overfit the data that CNN extracts and feeds into the LSTM model [26-30].

Figure 5 ; CNN model



Findings

Exploratory Data Analysis is beneficial for machine learning issues because it helps to increase the assurance that future outcomes will be reliable, correctly interpreted, and applicable to the required business settings. Only when the raw data has been confirmed and examined for abnormalities, confirming that the data set was accurately acquired, can such a level of certainty be attained. EDA aids in discovering ideas that 36 were not obvious or deserving of further study to researchers and business stakeholders. A data mining approach called Data Preprocessing entails putting raw data into a comprehensible format. Real-world data is frequently inaccurate and lacking in specific behaviours or trends. It is also frequently inconsistent and incomplete. The research was processed in the manner described below.

- Missing Values : It was discovered that there are few occurrences with null values, as per our EDA step. As a result, this becomes a crucial stage. We will categorise our instances according to the location and date, and then replace the null values with the corresponding mean values to impute the missing information. Feature Expansion:Date feature can be expanded to Day, Month and Year and then these newly created features can be further used for other preprocessing steps.

- Dummy variables : These are made artificially to represent attributes that have two or more separate categories or levels . However, because there are 16 different values, our single feature will now be changed into 16 different features, which leads to the curse of dimensionality. Each instance will have one feature with one value and the remaining fifteen features with zero values [31-33].

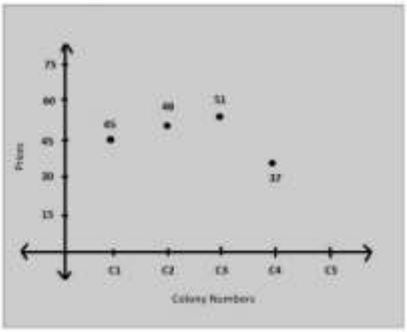
- Handling Class Imbalance: During the EDA process, we discovered that our data collection is incredibly unbalanced. Due to the fact that our model doesn't learn much about the minority class, biassed findings are produced by unbalanced data. We ran two tests, one using data that was oversampled and the other with data that was undersampled. \

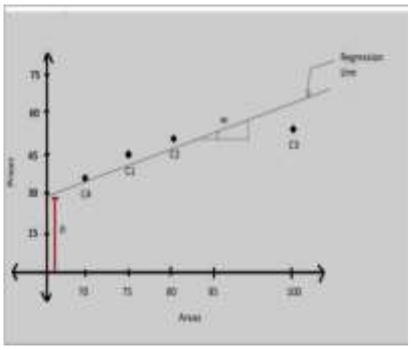
- Undersampling: To get rid of instances of the majority class, we used Imblearn's random under sampler library . In order to minimise information loss, this elimination is based on distance.

- Oversampling: To create synthetic instances for the minority class, we employed Imblearn's SMOTE approach. The minority class's subset of data is used as an example before additional artificial instances that are comparable to it are produced.

For modelling the link between a scalar output and one or more explanatory variables, it uses a linear approach (also known as dependent and independent variables). Simple linear regression is the situation where there is only one explanatory factor. In linear regression, linear predictor functions are used to model relationships, with the model's unknown parameters being estimated from the data. These models are referred to as linear models. The conditional mean of the response is typically considered to be a linear function of the values of the explanatory variables (or predictors); the conditional median or another quantile is occasionally employed. In common with all other types of regression analysis, linear regression concentrates on the conditional probability distribution of the response given the values of the predictors rather than the joint probability distribution of all these variables, which is the purview of multivariate analysis. Therefore, anytime we typically talk about machine learning, we are aware that there is data that has both features and labels. We then input this data into our machine in the hopes that it would predict the labels of fresh data based on its features. We can create a best fit line using a method called Simple Linear Regression, like in the prior scenario. However, this line would be more accurate if one additional detail had been supplied. Here, labels are referred to as dependent variables and characteristics as independent variables. Prices are the dependent variable in this case, whereas the independent variable is the area. The method we previously used to plot the graph will now be slightly altered. Here, the X-axis will now reflect the areas of the homes, while the Y-axis will show the prices of the homes. This will result in a 2D graph [34-36].

FIGURE 6; Plot of statistical data dependency





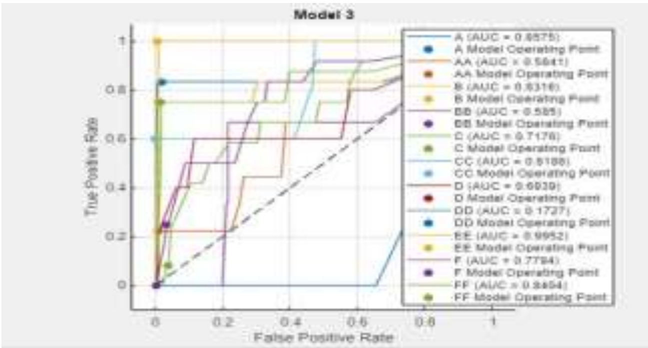
Result and discussion

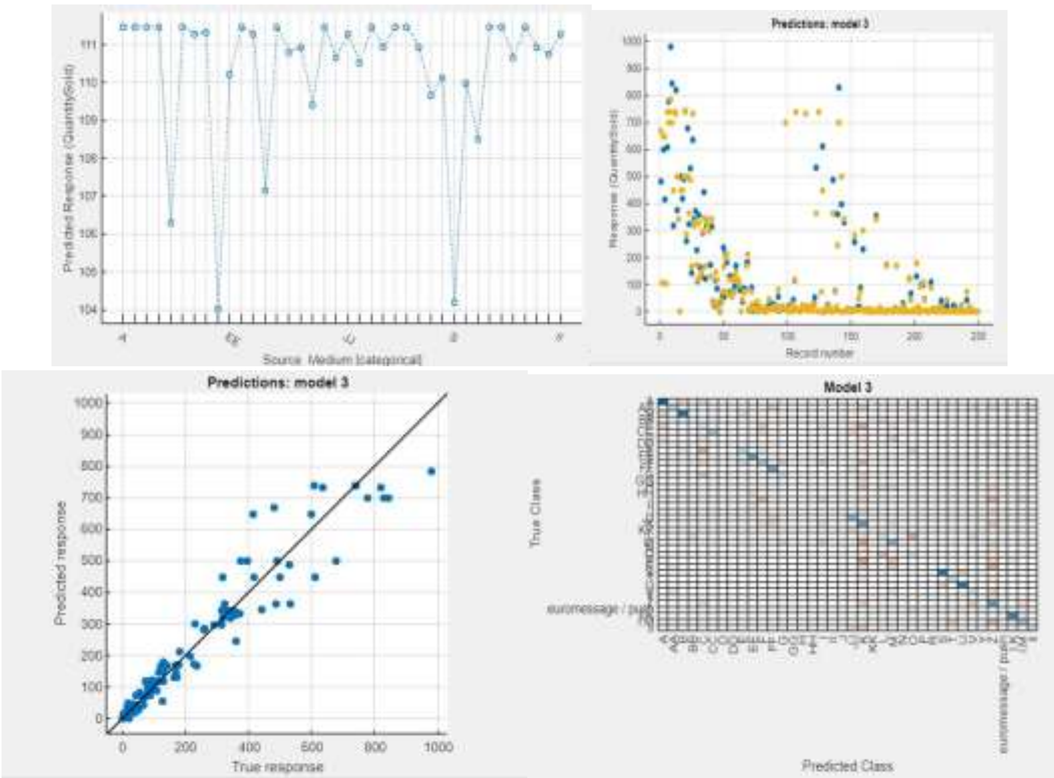
When rating classification models, the following metrics are reported. The proportion of accurate results to all instances is what constitutes accuracy. Aim for a high accuracy rate, accuracy is calculated as the number of accurate predictions divided by the number of data points. The classifier's advantage over a random prediction is measured by a single score called the log loss. By comparing the probability of your model's outputs to known values, the log loss gauges the uncertainty of your model (ground truth).. For the entire model, you want to reduce log loss as much as possible.Precision is the ratio of accurate outcomes to all positive findings. The proportion of all accurate outcomes that the model returns is known as recall.The F1-score is the weighted average of accuracy and recall between 0 and 1, with 1 being the optimal F-score.

Random splits are unavailable for some types of data. If we want to train a model for weather forecasting, for instance, we can't arbitrarily divide the data into training and testing sets. This will cause confusion in the seasonal pattern! Such information is typically referred to as "time series." In these situations, a time-wise split is used. The training set may consist of information from the last three years and the first ten months of the current year. During the last two months, the testing or validation set may be distributed. Additionally, the concept of window sets is used, in which the model is continually trained up to a specific date and then assessed on succeeding dates so that the training window keeps expanding [37-40].

One of the most effective methods for obtaining a stable model is Bootstrap. Considering that it adheres to the idea of random sampling, it is comparable to the random splitting approach. A sample size must be chosen in the first stage (which is usually equal to the size of the original dataset). After that, a sample data point from the initial dataset must be picked at random and added to the bootstrap sample. The sample has to be added to and then reinserted into the original sample. It is necessary to perform this procedure N times, where N is the sample size. Probabilistic measures consider both the complexity and the performance of the model. The degree to which a model can account for data variation is determined by its complexity. For instance, a neural network has a relatively high complexity level compared to a strongly biassed model like the linear regression technique. It's also vital to remember that only data from the training set is used to compute model performance for probabilistic metrics. Typically, a hold-out test set is not necessary The fact that probabilistic measurements do not take into account the models' uncertainty and have a tendency to favour simpler models over complicated models has certain drawbacks, though. Since the bootstrap sample is produced by sampling data points from the original dataset with replacement, it is a resampling approach. In light of this, the bootstrap sample can change in FIGURE 7 [41-44].

Figure 7; usability testing plots



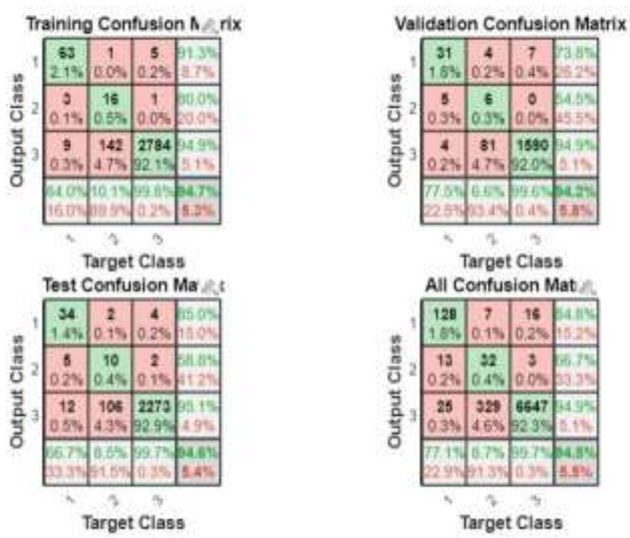


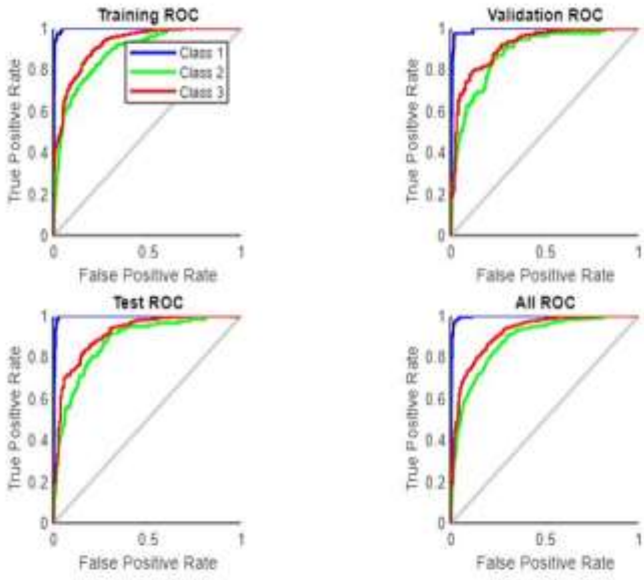
For instance, if we are finding frauds in bank data, the ratio of fraud to non-fraud occurrences can be 1:99. As a result, the model will shown to be 99% accurate by accurately predicting that no test case will contain fraud. The model, which is 99.9% accurate, will be totally useless. If a model is improperly trained and concludes that none of the 1000 (let's say) data points are fake, it will miss out on the 10 fraudulent data points. The model successfully predicted 990 data points, making its accuracy (990/1000)*100 = 99%, according to an evaluation of accuracy. Precision is thus a terrible indicator for the model's health.

We established two benchmarks for comparison: feature extraction and classification, in order to evaluate our detection methodology on malwares. On the same dataset, the experiment carries out three malware detection tasks. The findings are displayed in figure 8. The outcomes clearly show that Approach B's malware detection method surpassed Method A's and Method C's in terms of accuracy, recall, F1, and kappa. Particularly, Method A has attained 98% accuracy, which is particularly important for actual malware detection because misclassification has a large cost.

Since our approach relies on machine learning principles, we also employ the model evaluation method to examine the outcomes of our tests. For a more complete evaluation, we used the following main measuring metrics in this study: accuracy, precision, recall, F1, kappa, and ROC curve. Based on the correspondence between the actual class and the predicted class, we may divide the sample set into four groups in a binary classification task: true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) [45-46].Accuracy is determined by how much of the whole sample is made up of accurate samples. For a sample set D, the accuracy is given. precisely what is it? Recall is described thusly. Added performance indicators comprise. The receiver operating characteristic (ROC) curve is created by choosing the specificity (false-positive rate) and sensitivity (true-positive rate) as the horizontal and vertical axes, respectively. The corresponding area under the receiver operating characteristic (ROC) curve is one of the most used metrics for probabilistically evaluating the efficacy of classifiers (AUC) [47-48].

Figure 8; security testing plots





Conclusion

One of the most crucial aspects of e-commerce websites is security testing due to the handling of sensitive data and personal information. A security issue will unavoidably arise on an e-commerce website because of its enormous visibility, data, and information. the process of examining the security of an e-commerce website and the parts of it that hackers and other malicious users regularly target, such as the shopping cart, payment gateway, etc. An e-commerce website must go through comprehensive security testing to ensure that any data leak is prevented. Security testing should be performed by a team of certified testers with relevant industry experience. The information on the products they wish to buy is what visitors to e-commerce websites are seeking for. But only websites that can make this procedure as quick and easy as possible are of interest to them. They have a hard time with websites that are difficult or take a long time to load. Typically, it takes less than 10 to 15 seconds for users to leave a website that doesn't provide them anything of value. In a similar vein, most users will leave a website that takes more than three seconds to load in favour of a more useful alternative. As a result, an important aspect of interface design is ensuring sure a website is useful.Testing a website's usability only entails looking for outstanding usability. Testers should thus ensure that it keeps users on the website for a longer period of time. This overall paper emphasizes ion statistical feature selection based approach for usability testing of several systems and malware patch detection for security testing in e-commerce.

References

[1] M. J. Khalid *et al.*, “Integration of Discrete Wavelet Transform, DBSCAN, and Classifiers for Efficient Content Based Image Retrieval,” *Electronics*, vol. 9, no. 11, p. 1886, Nov. 2020, doi: 10.3390/electronics9111886.

[2] T. M. Lehmann *et al.*, “Automatic categorization of medical images for content-based retrieval and data mining,” *Computerized Medical Imaging and Graphics*, vol. 29, no. 2–3, pp. 143–155, Mar. 2005, doi: 10.1016/j.compmedimag.2004.09.010.

[3] A. Latif *et al.*, “Content-Based Image Retrieval and Feature Extraction: A Comprehensive Review,” *Mathematical Problems in Engineering*, vol. 2019, pp. 1–21, Aug. 2019, doi: 10.1155/2019/9658350.

[4] J. A. Miranda-Correa, M. K. Abadi, N. Sebe, and I. Patras, “AMIGOS: A Dataset for Affect, Personality and Mood Research on Individuals and Groups,” *arXiv:1702.02510 [cs, q-bio]*, Apr. 2017, Accessed: Nov. 30, 2021. [Online]. Available: http://arxiv.org/abs/1702.02510

[5] T. M. Dagnew and U. Castellani, “Advanced Content Based Image Retrieval for Fashion,” in *Image Analysis and Processing — ICIAP 2015*, vol. 9279, V. Murino and E. Puppo, Eds. Cham: Springer International Publishing, 2015, pp. 705–715. doi: 10.1007/978-3-319-23231-7_63.

[6] A. Babenko, A. Slesarev, A. Chigorin, and V. Lempitsky, “Neural Codes for Image Retrieval,” *arXiv:1404.1777 [cs]*, Jul. 2014, Accessed: Nov. 30, 2021. [Online]. Available: http://arxiv.org/abs/1404.1777

[7] M. Deshmane and S. Madhe, “ECG Based Biometric Human Identification Using Convolutional Neural Network in Smart Health Applications,” in *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, Pune, India, Aug. 2018, pp. 1–6. doi: 10.1109/ICCUBEA.2018.8697579.

[8] H. Feng, H. M. Golshan, and M. H. Mahoor, “A wavelet-based approach to emotion classification using EDA signals,” *Expert Systems with Applications*, vol. 112, pp. 77–86, Dec. 2018, doi: 10.1016/j.eswa.2018.06.014.

[9] Ka-Man Wong, Kwok-Wai Cheung, and Lai-Man Po, “MIRROR: An Interactive Content Based Image Retrieval System,” in *2005 IEEE International Symposium on Circuits and Systems*, Kobe, Japan, 2005, pp. 1541–1544. doi: 10.1109/ISCAS.2005.1464894.

[10] J. Laaksonen, M. Koskela, S. Laakso, and E. Oja, “PicSOM – content-based image retrieval with self-organizing maps,” *Pattern Recognition Letters*, vol. 21, no. 13–14, pp. 1199–1207, Dec. 2000, doi: 10.1016/S0167-8655(00)00082-9.

[11] M. Alojail and S. Bhatia, “A Novel Technique for Behavioral Analytics Using Ensemble Learning Algorithms in E-Commerce,” *IEEE Access*, vol. 8, pp. 150072–150080, 2020, doi: 10.1109/ACCESS.2020.3016419.

[12] P. Balaraman and S. Chandrasekar, “E-Commerce Trends and Future Analytics Tools,” *Indian Journal of Science and Technology*, vol. 9, no. 32, Aug. 216, doi: 10.17485/ijst/2016/v9i32/98653.

[13] J. K. Gerrikagoitia, I. Castander, F. Rebón, and A. Alzua-Sorzabal, “New Trends of Intelligent E-marketing Based on Web Mining for E-shops,” *Procedia - Social and Behavioral Sciences*, vol. 175, pp. 75–83, Feb. 2015, doi: 10.1016/j.sbspro.2015.01.1176.

[14] R. J. Kauffman, J. Srivastava, and J. Vayghan, “Business and data analytics: New innovations for the management of e-commerce,” *Electronic Commerce Research and Applications*, vol. 11, no. 2, pp. 85–88, Mar. 2012, doi: 10.1016/j.elerap.2012.01.001.

[15] Y. Li and N. Zhong, “Web mining model and its applications for information gathering,” *Knowledge-Based Systems*, vol. 17, no. 5–6, pp. 207–217, Aug. 2004, doi: 10.1016/j.knosys.2004.05.002.

[16] R. R. Nadikattu, “Research on Data Science, Data Analytics and Big Data,” *SSRN Journal*, 2020, doi: 10.2139/ssrn.3622844.

[17] U. Sutinen, H. Saarijärvi, and M. Yrjölä, “Shop at your own risk? Consumer activities in fashion e-commerce,” *Int J Consumer Studies*, vol. 46, no. 4, pp. 1299–1318, Jul. 2022, doi: 10.1111/ijcs.12759.

[18] C. Liu, S. Wang, and G. Jia, “Exploring E-Commerce Big Data and Customer-Perceived Value: An Empirical Study on Chinese Online Customers,” *Sustainability*, vol. 12, no. 20, p. 8649, Oct. 2020, doi: 10.3390/su12208649.

[19] C.-N. Wang, T.-T. Dang, N.-A.-T. Nguyen, and T.-T.-H. Le, “Supporting Better Decision-Making: A Combined Grey Model and Data Envelopment Analysis for Efficiency Evaluation in E-Commerce Marketplaces,” *Sustainability*, vol. 12, no. 24, p. 10385, Dec. 2020, doi: 10.3390/su122410385.

[20] R. Hasan, S. Palaniappan, S. Mahmood, A. Abbas, K. U. Sarker, and M. U. Sattar, “Predicting Student Performance in Higher Educational Institutions Using Video Learning Analytics and Data Mining Techniques,” *Applied Sciences*, vol. 10, no. 11, p. 3894, Jun. 2020, doi: 10.3390/app10113894.

[21] “UNIT IV ASSOCIATION RULE MINING AND CLASSIFICATION,” p. 37.

[22] H. Yadav, U. Soni, S. Gupta, and G. Kumar, “Evaluation of Barriers in the Adoption of E-Commerce Technology in SMEs: A Fuzzy DEMATEL Approach,” *Journal of Electronic Commerce in Organizations*, vol. 20, no. 1, pp. 1–18, Jan. 2022, doi: 10.4018/JECO.292472.

[23] A. Razaque, M. B. H. Frej, B. Alotaibi, and M. Alotaibi, “Privacy Preservation Models for Third-Party Auditor over Cloud Computing: A Survey,” *Electronics*, vol. 10, no. 21, p. 2721, Nov. 2021, doi: 10.3390/electronics10212721.

[24] P. Patil, P. Narayankar, Narayan D.G., and Meena S.M., “A Comprehensive Evaluation of Cryptographic Algorithms: DES, 3DES, AES, RSA and Blowfish,” *Procedia Computer Science*, vol. 78, pp. 617–624, 2016, doi: 10.1016/j.procs.2016.02.108.

[25] P. Patil, P. Narayankar, Narayan D.G., and Meena S.M., “A Comprehensive Evaluation of Cryptographic Algorithms: DES, 3DES, AES, RSA and Blowfish,” *Procedia Computer Science*, vol. 78, pp. 617–624, 2016, doi: 10.1016/j.procs.2016.02.108.

[26] E. Idzikowska and K. Bucholc, “Error Detection Schemes for CED in Block Ciphers,” in *2008 IEEE/IFIP International Conference on Embedded and Ubiquitous Computing*, Shanghai, China, Dec. 2008, pp. 22–27. doi: 10.1109/EUC.2008.134.

[27] N. Joshi, K. Wu, and R. Karri, “Concurrent Error Detection Schemes for Involution Ciphers,” in *Cryptographic Hardware and Embedded Systems - CHES 2004*, vol. 3156, M. Joye and J.-J. Quisquater, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 400–412. doi: 10.1007/978-3-540-28632-5_29.

[28] X. Zhang and K. K. Parhi, “On the Optimum Constructions of Composite Field for the AES Algorithm,” *IEEE Trans. Circuits Syst. II*, vol. 53, no. 10, pp. 1153–1157, Oct. 2006, doi: 10.1109/TCSII.2006.882217.

[29] Q. Zheng, N. Liu, and F. Wang, “An Adaptive Embedding Strength Watermarking Algorithm Based on Shearlets’ Capture Directional Features,” *Mathematics*, vol. 8, no. 8, p. 1377, Aug. 2020, doi: 10.3390/math8081377.

[30] P. Visconti, S. Capoccia, E. Venere, R. Velázquez, and R. de Fazio, “10 Clock-Periods Pipelined Implementation of AES-128 Encryption-Decryption Algorithm up to 28 Gbit/s Real Throughput by Xilinx Zynq UltraScale+ MPSoC ZCU102 Platform,” *Electronics*, vol. 9, no. 10, p. 1665, Oct. 2020, doi: 10.3390/electronics9101665.

[31] R. Saha, G. Geetha, G. Kumar, and T. Kim, “RK-AES: An Improved Version of AES Using a New Key Generation Process with Random Keys,” *Security and Communication Networks*, vol. 2018, pp. 1–11, Nov. 2018, doi: 10.1155/2018/9802475.

[32] M. I. S. Reddy and A. P. S. Kumar, “Secured Data Transmission Using Wavelet Based Steganography and Cryptography by Using AES Algorithm,” *Procedia Computer Science*, vol. 85, pp. 62–69, 2016, doi: 10.1016/j.procs.2016.05.177.

[33] M. Feldhofer, S. Dominikus, and J. Wolkerstorfer, “Strong Authentication for RFID Systems Using the AES Algorithm,” in *Cryptographic Hardware and Embedded Systems - CHES 2004*, vol. 3156, M. Joye and J.-J. Quisquater, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 357–370. doi: 10.1007/978-3-540-28632-5_26.

[34] N. Mathur and R. Bansode, “AES Based Text Encryption Using 12 Rounds with Dynamic Key Selection,” *Procedia Computer Science*, vol. 79, pp. 1036–1043, 2016, doi: 10.1016/j.procs.2016.03.131.

[35] E. Oswald, S. Mangard, N. Pramstaller, and V. Rijmen, “A Side-Channel Analysis Resistant Description of the AES S-Box,” in *Fast Software Encryption*, vol. 3557, H. Gilbert and H. Handschuh, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 413–423. doi: 10.1007/11502760_28.

[36] R. Damaševičius and L. Zailskaitė-Jakštė, “Usability and Security Testing of Online Links: A Framework for Click-Through Rate Prediction Using Deep Learning,” *Electronics*, vol. 11, no. 3, p. 400, Jan. 2022, doi: 10.3390/electronics11030400.

[37] K. Diamantaras, M. Salampasis, A. Katsalis, and K. Christantonis, “Predicting Shopping Intent of e-Commerce Users using LSTM Recurrent Neural Networks:,” in *Proceedings of the 10th International Conference on Data Science, Technology and Applications*, Online Streaming, --- Select a Country ---, 2021, pp. 252–259. doi: 10.5220/0010554102520259.

[38] I. Dumanska, L. Hrytsyna, O. Kharun, and O. Matviiets, “E-commerce and M-commerce as Global Trends of International Trade Caused by the Covid-19 Pandemic,” *WSEAS TRANSACTIONS ON ENVIRONMENT AND DEVELOPMENT*, vol. 17, pp. 386–397, Apr. 2021, doi: 10.37394/232015.2021.17.38.

[39] S. Kandula, S. Krishnamoorthy, and D. Roy, “A prescriptive analytics framework for efficient E-commerce order delivery,” *Decision Support Systems*, vol. 147, p. 113584, Aug. 2021, doi: 10.1016/j.dss.2021.113584.

[40] D. Koehn, S. Lessmann, and M. Schaal, “Predicting online shopping behaviour from clickstream data using deep learning,” *Expert Systems with Applications*, vol. 150, p. 113342, Jul. 2020, doi: 10.1016/j.eswa.2020.113342.

[41] F. Li and Y. Li, “Usability evaluation of e-commerce on B2C websites in China,” *Procedia Engineering*, vol. 15, pp. 5299–5304, 2011, doi: 10.1016/j.proeng.2011.08.982.

[41] V. L. F. Minatogawa *et al.*, “Operationalizing Business Model Innovation through Big Data Analytics for Sustainable Organizations,” *Sustainability*, vol. 12, no. 1, p. 277, Dec. 2019, doi: 10.3390/su12010277.

[43] Y. Ozyurt, T. Hatt, C. Zhang, and S. Feuerriegel, “A Deep Markov Model for Clickstream Analytics in Online Shopping,” in *Proceedings of the ACM Web Conference 2022*, Virtual Event, Lyon France, Apr. 2022, pp. 3071–3081. doi: 10.1145/3485447.3512027.

[44] A. S. Rajawat, K. Barhanpurkar, S. B. Goyal, P. Bedi, R. N. Shaw, and A. Ghosh, “Efficient Deep Learning for Reforming Authentic Content Searching on Big Data,” in *Advanced Computing and Intelligent Technologies*, vol. 218, M. Bianchini, V. Piuri, S. Das, and R. N. Shaw, Eds. Singapore: Springer Singapore, 2022, pp. 319–327. doi: 10.1007/978-981-16-2164-2_26.

[45] P. N. M. Rizki, I. Handoko, P. Purnama, and D. Rustam, “Promoting Self-Regulated Learning for Students in Underdeveloped Areas: The Case of Indonesia Nationwide Online-Learning Program,” *Sustainability*, vol. 14, no. 7, p. 4075, Mar. 2022, doi: 10.3390/su14074075.

[46] C. C. Saw and A. Inthiran, “Designing for Trust on E-Commerce Websites Using Two of the Big Five Personality Traits,” *JTAER*, vol. 17, no. 2, pp. 375–393, Mar. 2022, doi: 10.3390/jtaer17020020.

[47] W. Serrano, “Neural Networks in Big Data and Web Search,” *Data*, vol. 4, no. 1, p. 7, Dec. 2018, doi: 10.3390/data4010007.

[48] B. Sohrabi, P. Mahmoudian, and I. Raeesi, “A framework for improving e-commerce websites usability using a hybrid genetic algorithm and neural network system,” *Neural Comput & Applic*, vol. 21, no. 5, pp. 1017–1029, Jul. 2012, doi: 10.1007/s00521-011-0674-7.