I.J. Information Engineering and Electronic Business, **2023, 5, 1-12**
Published Online on October 8, 2023 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijieeb.2023.05.01

Exploration on Quick Response (QR) Code Behaviour in Commerce based Platforms Using Machine Learning

Archana Uriti*
GMR Institute of Technology, Rajam, Andhra Pradesh, India
Email: archana.u@gmrit.edu.in
ORCID ID: https://orcid.org/0000-0001-9123-3296
*Corresponding Author

Surya Prakash Yalla
GMR Institute of Technology, Rajam, Andhra Pradesh, India
Email : suryaprakash.y@gmrit.edu.in
ORCID ID: https://orcid.org/0000-0002-0896-0714

Received: 15 March, 2023; Revised: 10 April, 2023; Accepted: 04 May, 2023; Published: 08 October, 2023

Abstract: The "rapid response" code, or QR code, is made to quickly decode vast amounts of data. Any managed device, such as a smartphone, is able to capture it, and it is simple to access simply scanning the 2D matrix code. The dataset is analyzed utilizing machine learning techniques, such as the confusion matrix score utilized for the multinomial naive Bayes algorithm's performance analysis. The QR code generation is limited to single product and is extended now to include all products. Due to its ability to provide clients with benefits including speedy, error-free access and the ability to store a lot of data. Generally, many people are using the online payment for any transaction for flexibility and one can do at any place at any time. For bulk or huge payment, cash is not a good option. Hence many retailers join in the e-wallet companies and make their payment so flexible and faster transaction. Because of these benefits, QR code has becoming widespread.

Index Terms: Confusion matrix, QR code, Count vectorizer, GridsearchCV, Python Tkinter.

1. Introduction

Nowadays the QR code popularity is increasing rapidly by the smartphone users. The existing QR code is limited to few products accessing and for payment. The idea of generating QR code is taken to overcome with the limitation of any Ecommerce marketing. To access any products, the QR is limited to specific product and one cannot access all products in one specific zone. The proposed methodology is to generate QR code with a benefit of including all products captured at one click and gives the speedy and error free access. With a white background and a square grid of black squares, a QR code[1] can be read by a camera or other image-reading equipment. Here, textual data is transformed into vectors and processed using the TF-IDF vectorizer in order to analyse the product data. After the data has been successfully transformed into vector form, these vectors are sent into the classification algorithm[2], which outputs the product category when provided the input as a description. By linking several identities in a single domain while obscuring information, the QR code modernizes the process of identifying products or people.

The remaining paper is organized in such a way that, Section II gives the information of related work, and Section III shows the proposed work. The results are discussed in Section IV. The paper finally comes to an end in Section V.

2. Related Works

The articles that followed used a range of techniques, including the naive Bayes algorithm, count vectorizer, TF-IDF, and also Grid search CV, which is used for fine-tuning the parameters and providing a classification report. The authors M. Bala Krishna, et al., presented their mobile app to fight counterfeiting in their article Using QR Codes to Authenticate Products. This study discusses the detection of QR for a single product's information piece along with authentication. Safe and secure transactions are the main uses of QR. Large amounts of data that can be expressed on the horizontal and vertical axes are contained in the encrypted QR. Depending on the type of encoded data, a QR code's

capacity may vary. And error correction level, The REED-Solomon algorithm[3] is used here i.e; forward error correction this algorithm helps in recovering the corrupted messages by using the encoder and decoder. The encoder portion of this technique adds the parity bits to the original bits, while the decoder portion decodes that data and corrects the problem by identifying the corrupted messages. In the suggested method, a single product's QR code is generated. Here, the characteristics of the product are utilised to ascertain its authenticity, and a server is kept up to comprehensively update the product list.

Iuliia Tkachenko, William Puech, Senior member IEEE, Christophe Destruel, Olivier Strauss, Jean-Marc Gaudin, and Christian Guichard developed a two-level QR code that allows for the transmission of private conversations and the authentication of documents.The strategy suggested. The quick response code project is mostly used to store certain data and run fast apps. In this paper, document validation will often be done using QR's two-level storage layers. For those who know how, a QR code is encoded in binary format. Each byte of information is represented by a module that is either black or white. The typical QR code has the public message Mpub embedded within it using the traditional creation method. The standard procedure for creating QR codes includes the stages listed below. The best mode (alphanumeric or numeric) is first chosen by reviewing the message content. The main use of the 2LQR code is to change the black modules to textured patterns related to the codeword C private. Replace the black Modules with columns that go from bottom to top in the appropriate corner of the QR code[2].

The acronym "QR" stands for "Quick Response." Denso Wave Corporation in Japan created it. "Quick Response" is denoted by the abbreviation "QR." It was produced by Japan's Denso Wave Corporation. More data can be encoded in a QR code, a two-dimensional barcode, than in a one-dimensional barcode. These benefits have led to a widespread adoption of QR codes throughout the world. The study makes an effort to emphasise QR codes, their benefits, advantages, and shortcomings. It is necessary for the scanning The use of QR codes is not universal. Therefore, not everyone who sees a QR code symbol will take out his phone and take a picture of it. Not everyone has a camera phone, and many smartphones do not have a QR scanner. Additionally, customers may be directed to a website that displays incorrectly on a mobile device.. The written code on the product cannot be scanned if it is completely broken during transportation[3].

Text categorization using Bayesian Naive Bayes classifiers the authors of this article outlined The nave Bayes hypothesis claims that the Bayes theorem underlies the collection of classification algorithms. One technique for probabilistic learning is the Multinomial Naive Bayes algorithm. The premise that each feature being classified is unique from the other characteristics forms the foundation of a naive Bayes classifier, which combines many techniques. The presence or absence of one trait does not affect the presence or absence of the opposite trait. The Bernoulli Naive Bayes method, one of the Naive Bayes algorithms used in machine learning, is great for use in binary distributions where the output label may be absent or present. It is best utilised when the output label is either present or missing and the dataset has a fairly binary distribution. This algorithm's greatest benefit is that it only accepts[4].

This essay makes it very clear that Python provides a number of options for developing graphic computer programmes (GUI). When compared to any or all Graphical user interface techniques, the method that leverages Tkinter most successfully. It uses the typical Python interface of the Tk GUI toolkit[5]. The architecture of this layer is the graphical user interface. The development of graphical user interfaces for scripts that automate the experimenting process is the primary goal of this layer. The main packages have been implemented and are being used right now. Tkinter modules The Tkinter module offers event loops for the receiving of user events such as button pushes, keystrokes, mouse movements, and window resizing in contrast to operating system-controlled window elements.

This essay clearly demonstrates how Python offers a variety of alternatives for creating graphic computer programmes (GUI). The approach that best utilises Tkinter when compared to any or all Graphical user interface approaches. It makes use of the Tk GUI toolkit's standard Python interface. The graphical user interface serves as this layer's architecture. This layer's main objective is to create graphical user interfaces for scripts that automate the experimentation process. The primary packages have been set up and are currently in use. modules for Tkinter In contrast to operating system-controlled window elements, the Tkinter module provides event loops for the reception of user events including button presses, keystrokes, mouse movements, and window resizing. MNB: Multinomial Naive Bayes [6]and SVM accuracy for the provided dataset is 63.02 percent. The model will assist application engineers in classifying automatic fault detection. This approach will improve the accuracy and effectiveness of the defect classification process. This model can be employed with any text dataset and applied to any issue that is comparable.

Emotional assessment is one of the most common NLP tasks[7], it has been found. These days, there is a lot of interest in emotional analysis. According to the study's findings, the hardest problems are with client categorization and emotional growth. Sentence-level classification seeks to organise various positive or negative phrases based on the sentiment they convey. High vectors come from reviews that have at least a 4-star rating, whilst negative vectors come from reviews with 1- and 2-star ratings. Develop neutral class vectors in accordance with 3-star ratings. The study of people's attitudes and feelings toward various objects is known as sentiment analysis[8], also referred to as opinion mining. The essential topic of emotion is covered in this article.

The authors of this essay provide comprehensive information on utilising HTML and CSS to construct websites. Responsive web design aims to produce websites that offer the best user experience possible. Responsive web design's [9] main principle is to scale content to fit the size of the screen without hiding any material or changing how the layout looks. Building websites using flexible code has become crucial since more and more people are utilising portable equipment to view websites. Customers can readily access material because websites are now responsive to customers'

needs and the devices they are using. The size and capacity of the gadget influence how the site is structured. A modern website may help any company increase its visibility to potential customers. Websites are often used by businesses, institutions, organisations[10], and individuals for audience and customer engagement. Users may enter data into gadgets with a mouse or their fingertips, and there are various web browsers, screen sizes, and resolutions. There are also distinctions between the systems and browsers. Responsive web design is not "one size fits all" for all devices.

Sentiment is a method for determining out whether a paragraph of literature is uplifting, depressing, or neutral. A very useful tool for expressing the views of the general public, a group, or an individual is data sentiment analysis. Here, movie reviews are analysed using naive Bayes, k-nearest neighbour, and random forest techniques. Data cleansing must be done first to reduce duplication and eliminate noise. After that, the token is given and preserved for analysis, and an algorithm is used. The algorithms used were Random Forest, K-Nearest Neighbor, and Naive Bayes[11]. The Nave Bayes classifier produced the best results. The Naive Bayes classifier had an accuracy of 81.45%, the Random Forest classifier had an accuracy of 78.65%, and the K-Nearest Neighbor classifier had an accuracy of 55.30%. In today's environment, communication is crucial to day-to-day living. Since the dawn of time, people have been able to communicate with one another through a variety of means.

In today's environment, communication is crucial to day-to-day living. Since the dawn of time, people have been able to communicate with one another through a variety of means. Developers of software that must store and make available user-generated data, such as images or other types of data, were the target audience for the development of Firebase Storage. It provides safe file transfers and downloads for Firebase apps[12] regardless of the network's status. Firebase Storage is available through Google Cloud Storage, an expensive, dependable, and user-friendly object storage server. Many developers are presently using web services like Firebase, which provide real-time database connectivity, to construct messaging applications. Given that it provides developers with a comprehensive solution, Firebase is a step in the right direction for application development. Because more businesses will adopt it, application development will be significantly quicker and more effective.

3. Proposed Work

The proposed work is divided into the following six modules keeping every module consisting of main functionality in the whole work that is showed in the below Fig.1.

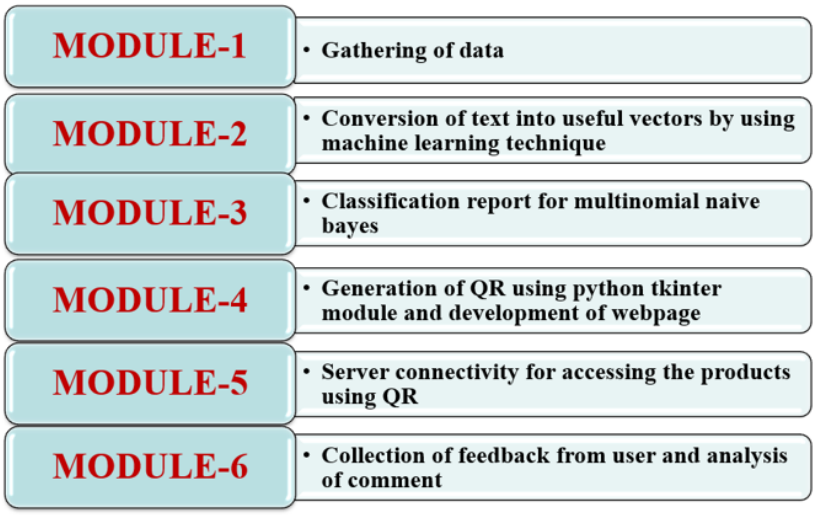

Fig. 1. Flow diagram of proposed work

3.1 Gathering of data:

Here the products dataset is taken from the outsource and the data is imported using commands. The natural language tool kit is imported and the stop words from corpus are imported which helps to eliminate the unimportant words and helps to focus on useful words. Now the dataset is analysed and splitted into 3 categories:

1) Main Category
2) Sub Category
3) Sub- Sub Category

The below Fig.2 shows some of the products based on the categories and the Fig.3 represents the quantity of the categories.

product_name	product_category_tree	retail_price	discounted_price
Alisha Solid Women's Cycling Shorts	Clothing >> Women's Clothing >> Lingerie, Slee...	999.0	379.0
FabHomeDecor Fabric Double Sofa Bed	Furniture >> Living Room Furniture >> Sofa Bed...	32157.0	22646.0
AW Bellies	Footwear >> Women's Footwear >> Ballerinas >> ...	999.0	499.0
Alisha Solid Women's Cycling Shorts	Clothing >> Women's Clothing >> Lingerie, Slee...	699.0	267.0
Sicons All Purpose Arnica Dog Shampoo	Pet Supplies >> Grooming >> Skin & Coat Care >...	220.0	210.0

Fig. 2. Products based on categories

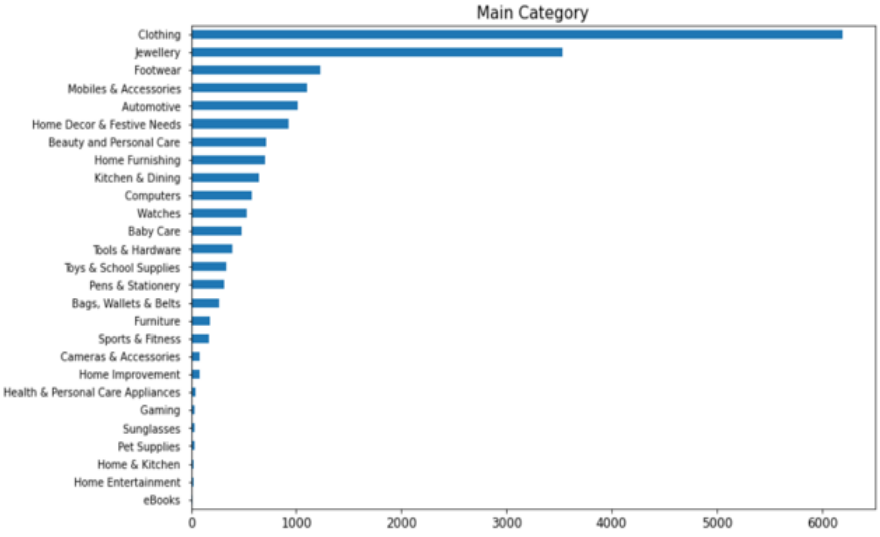

Fig. 3. Products quantity based on categories

3.2 Conversion of text into useful vectors by using machine learning technique

Here the model training is done by taking description and category of products.Here the CountVectorizer pipeline[13] and TFIDvectorizer pipeline are shown i.e based on the no. of words that occurs in the textTerm frequency inverse document frequency, the count vectorizer is transforming the text into vectors count vectorizer is for transforming the text into vectors. This TFIDF has two concepts:

- Term frequency : Number of occurance of specific term in the document.
- Inverse Document frequency : it shows the term is common on rare in the corpus or the measure of term.

Crossvalidation is employed in this case together with the gridsearchCV technique. In this case, the GridsearchCV approach is utilised to find the optimum parameter values from the grid of parameters that has been provided. In order to evaluate the model for each combination of values provided into the dictionary, GridsearchCV takes into account all possible value combinations. The resulting accuracy or loss for each combination of hyperparameters allows one to select the optimum parameter[14]. GridSearch and cross-validation[15] both require a long time to analyse the optimum hyperparameters collectively.

Here the GridsearchCV method can be initialized as:

clf =GridsearchcCV(estimator,param_grid,cv,scoring) Those 4 arguments are:

1. Estimator : Using Scikit-learn

2. Param_grid : a dictionary with keys consisting of a list of parameter names and values.
3. Scoring : The gauge of performance. Precision is useful for classification models, for instance.
4. cv : A number that represents how many folds there are in a K-fold cross-validation.

GridSearchCV will give the best values for the specific hyperparameters. The model performance mainly depends on the value of hyperparameters. Here trying all the possible values to know the optimal value because there is no chance to get the best value for hyperparameters. GridSearchCV will consume time and resources when used to automate hyperparameter tweaking. GridSearchCV is a model selection function in the SK Learn package. For this to work, the Scikit Learn library must be installed on the Computer. By looping through provided hyperparameters, this function assists in fitting your estimator (model) to your training set. Select the best parameters from the list of hyperparameters at the end.

3.3 Training and Testing scores:

Machine Learning builds the predictive models based on the data which is divided into two sets, the first one is known as training data. In the trained data set, again it can be classified as labelled and unlabelled. By considering the data set, any machine learning algorithm can be build and generally, Training data set is larger in size than testing data. When the machine learning algorithm receives the data, it will analyse the precedent to develop an appropriate model. It will increase the outcome's precision. When dealing with massive amounts of data, machine learning algorithms will provide accurate predictions.

- Test the data when the model is built with trained data. By using AI, optimize and adjust our results can be possible for improving the data.
- Testing data is used to evaluate the model's correctness, whereas training data is a smaller subset of the original data.
- Trained and test datasets splitting ratio are 80:20, 70:30 or 90:10.

3.4 Classification Report :

It is one of the metrics used to assess the performance of a classification-based machine learning model. It indicates your model's precision, recall, F1 score, and support. That helps determining the overall effectiveness of our trained model easy. You must be conversant with each and every metric displayed in a machine learning model's categorization report. The classification report of your machine learning model shows the accuracy, recall, F1 Score, and support score of your trained classification model, which is a performance assessment statistic in machine learning. The TFIDF vectorizer has a higher accuracy score when applying confusion matrix metrics to analyse the classification report. As a result, TF-IDF outperforms Count Vectorizer because it considers both the relevance of the words and their frequency in the corpus. Thus, by reducing the size of the input space, we may speed up the model-building process by removing words that aren't necessary for analysis. These accuracy levels are thus utilised to categorise things in the following modules that is shown in below Fig.4.

Metrics	Definition
Precision	Precision is defined as the ratio of true positives to the sum of true and false positives.
Recall	Recall is defined as the ratio of true positives to the sum of true positives and false negatives.
F1 Score	The F1 is the weighted harmonic mean of precision and recall. The closer the value of the F1 score is to 1.0, the better the expected performance of the model is.
Support	Support is the number of actual occurrences of the class in the dataset. It doesn't vary between models, it just diagnoses the performance evaluation process.

Fig. 4. classification metrics

3.5 Generation of QR using Tkinter module and development of web page

Here, Python tools are used to generate the QR code. Here, the Tkinter tool is primarily used. A window is made and labelled for the message box when it is imported from the Tkinter module. Now, a button has been added and a function for creating QR codes in Python has been built. So, following the successful generation of the QR shown in below Fig.5 and will get a dialogue window stating that the QR was successfully generated. As a result, the QR code is visible, and the website has been created to give customers a clearer picture of how to access things.

```
def generateCode():
    #Creating a QRCode object of the size specified by the user
    qr = qrcode.QRCode(version = size.get(),
                       box_size = 10,
                       border = 5)
    qr.add_data(text.get()) #Adding the data to be encoded to the QRCode object
    qr.make(fit = True) #Making the entire QR Code space utilized
    img = qr.make_image() #Generating the QR Code
    fileDirec=loc.get()+'\\'+name.get() #Getting the directory where the file has to be save
    img.save(f'{fileDirec}.png') #Saving the QR Code
    #Showing the pop up message on saving the file
    messagebox.showinfo("DataFlair QR Code Generator","QR Code is saved successfully!")

#Label for the window
headingFrame = Frame(wn,bg="azure",bd=5)
headingFrame.place(relx=0.15,rely=0.05,relwidth=0.7,relheight=0.1)
headingLabel = Label(headingFrame, text="Generate QR Code with DataFlair", bg='azure', font=('Times',20,'bold'))
headingLabel.place(relx=0,rely=0, relwidth=1, relheight=1)
```

Fig. 5. QR Generation

Now, HTML, CSS, and JS are being used to build the webpage. Cascading Style Sheets (CSS) documents can be included in an HTML document to describe how material is presented on a web page. A web page may also include JavaScript or Web Assembly applications, which are executed by the web browser, to provide it dynamic behaviour. Here, the main page's header has links to other pages, such the store page. Customers may quickly access the products on the store page and add them to their shopping carts. They have the option to modify the quantity while adding the products. Then, this website is connected to the QR code via Google's Firebase server. The header, footer, homepage, and store page make up this website. Before purchasing the products, users can alter the quantity of the items in their cart. As a consequence, consumers will benefit from the ability to purchase many items, as well as the fact that it is free to use and has an easy-to-use graphical user interface. A web page's URI (uniform resource locator) is unique (URL). When a user enters the URL of a web page into the browser, it downloads an HTML file from a web server and converts all of its contents into an interactive visual representation on the user's device. If the user clicks, taps, or otherwise activates the hyperlink, the browser repeats the process to load the page to which the hyperlink goes, which may be on the same or a different website.

3.6 Server connectivity for accessing the products using QR:

Here these webpages are deployed in the Firebase which is the hosting party of our webpage. Here the webpage is deployed and the link is generated after the successful deployment of webpage. The below Fig.6 shown the firebase console.

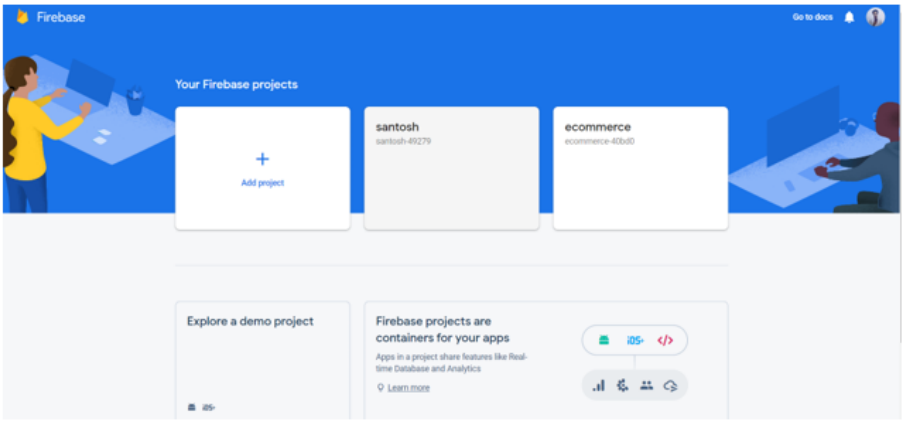

Fig. 6. Firebase console

After deploying the site the link is attached to the QR that is shown in below Fig.7 and now it is free for accessing the product by simply scanning the QR.

Fig. 7. Final QR for E-Commerce

3.7 Collection of feedback from the user and Analysis of comment:

After the feedback is submitted by the customer based on their experience and the analysis is done on the feedback using sentimental analysis. The algorithm estimates whether a review will be good or negative based on the comment. As a result, user input is gathered, and rating categories are created depending on the feedback provided by the user and their level of satisfaction that is shown in below Fig.8.

	product title	Customer name	Rating	comment	Category	typeofreview	text_final
0	red tshirt	Rishikumar Thakur	5	[product, is, very, good]	tshirt	1	['product', 'good']
1	red tshirt	Raza ji	4	[fits, good]	tshirt	1	['fit', 'good']
2	hrx shoes	Vaibhav Patel	4	[material, is, excellent]	shoes	1	['material', 'excellent']
3	hrx trouser	ishanth rajput	4	[fantastic]	trouser	1	['fantastic']
4	hrx shoes	sundar tehar	5	[good, fit]	shoes	1	['good', 'fit']
5	hrx trouser	Arun Yajurvedi	3	[good, quality, of, product]	trouser	1	['good', 'quality', 'product']
6	hrx shoes	Irukulla bharath	3	[good, fit]	shoes	1	['good', 'fit']
7	hrx socks	rithwik	4	[nice, quality]	socks	1	['nice', 'quality']
8	hrx trouser	siraj	5	[worthy]	trouser	1	['worthy']
9	hrx trouser	Sharad	5	[good]	trouser	1	['good']
10	hrx socks	Rishikumar Thakur	4	[elastic]	socks	1	['elastic']
11	hrx socks	Raza ji	3	[good]	socks	1	['good']
12	hrx socks	Vaibhav Patel	5	[better, quality]	socks	1	['good', 'quality']
13	black watch	rahul	4	[good, looking]	watch	1	['good', 'look']
14	red tshirt	hari	5	[product, is, very, good]	tshirt	1	['product', 'good']
15	red tshirt	shankar	4	[fits, good]	tshirt	1	['fit', 'good']
16	hrx shoes	reddy	4	[material, is, excellent]	shoes	1	['material', 'excellent']
17	hrx trouser	hemanth	4	[fantastic]	trouser	1	['fantastic']
18	hrx shoes	sunil	5	[good, fit]	shoes	1	['good', 'fit']
19	hrx trouser	shetty	1	[bad]	trouser	0	['bad']
20	hrx shoes	vicky	3	[good, fit]	shoes	1	['good', 'fit']
21	hrx socks	eswar	2	[too, costly, socks]	socks	0	['costly', 'sock']

Fig. 8. Division of rating for sentimental analysis

4. Results

4.1 Training data and Testing Data:

When the model has been suitably trained using the relevant training data, it is tested using test data. The complete training and testing method may be divided into three steps:

1. Feed: Before we can train the model, we must first provide it with training data.
2. Define: With supervised learning, the training data is now tagged with the required outputs, and the model translates the training data into text vectors or a variety of data attributes.
3. Test: Finally, we put the model to the test by giving it either test data or an unknown dataset. This stage ensures that the model has been adequately trained and has strong generalisation capabilities.

Training and Testing scores of Count Vectorizer and TFID vectorizer are given as below Fig.9.

```
print(clf.score(X_test, y_test))# Scoring Training data on CountVectorizer
print(clf.score(X_train, y_train))# Scoring Testing data on CountVectorizer
print(clf2.score(X_test, y_test))# Scoring Training data on TFIDFVectorizer
print(clf2.score(X_train, y_train))# Scoring Testing data on TFIDFVectorizer
```

```
0.9682539682539683
0.9939615984802226
0.9763939763939764
0.997896736549291
```

Fig. 9. Training and testing scores for TFID and CountVectorizer

4.2 Classification Report:

It is one of the metrics used to evaluate the effectiveness of classification-based machine learning models. Your model's accuracy, recall, F1 score, and support are displayed. It provides us with a more complete picture of our trained model's overall performance. To interpret a machine learning model's categorization report, you must be familiar with each statistic displayed. The below Fig.10 shows the classification report for count vectorizer and Fig.11 shows the classification report for TFIDF vectorizer.

	precision	recall	f1-score	support
Automotive	0.99	0.99	0.99	229
Baby Care	0.88	0.85	0.87	128
Bags, Wallets & Belts	0.93	0.75	0.83	71
Beauty and Personal Care	0.97	0.99	0.98	168
Cameras & Accessories	0.72	0.95	0.82	22
Clothing	0.99	0.99	0.99	1578
Computers	0.97	0.90	0.93	153
Footwear	0.99	0.99	0.99	309
Furniture	1.00	0.98	0.99	45
Gaming	0.54	0.88	0.67	8
Health & Personal Care Appliances	1.00	0.82	0.90	11
Home & Kitchen	0.57	1.00	0.73	8
Home Decor & Festive Needs	0.97	0.99	0.98	258
Home Entertainment	0.50	0.50	0.50	4
Home Furnishing	0.96	1.00	0.98	179
Home Improvement	0.71	0.85	0.77	20
Jewellery	0.99	1.00	0.99	875
Kitchen & Dining	0.98	0.96	0.97	141
Mobiles & Accessories	0.99	0.94	0.97	287
Pens & Stationery	0.87	0.73	0.79	74
Pet Supplies	0.80	1.00	0.89	4
Sports & Fitness	0.80	0.91	0.85	35
Sunglasses	1.00	1.00	1.00	11
Tools & Hardware	0.92	0.94	0.93	104
Toys & School Supplies	0.71	0.90	0.80	71
Watches	1.00	0.98	0.99	118
eBooks	1.00	0.33	0.50	3
accuracy			0.97	4914
macro avg	0.88	0.89	0.87	4914
weighted avg	0.97	0.97	0.97	4914

Fig. 10. Classification report for count vectorizer

	precision	recall	f1-score	support
Automotive	0.98	1.00	0.99	229
Baby Care	0.97	0.82	0.89	128
Bags, Wallets & Belts	0.94	0.86	0.90	71
Beauty and Personal Care	0.96	0.99	0.98	168
Cameras & Accessories	1.00	0.91	0.95	22
Clothing	0.99	1.00	0.99	1578
Computers	0.92	0.95	0.94	153
Footwear	0.99	0.99	0.99	309
Furniture	0.98	0.98	0.98	45
Gaming	0.88	0.88	0.88	8
Health & Personal Care Appliances	1.00	0.82	0.90	11
Home & Kitchen	1.00	0.75	0.86	8
Home Decor & Festive Needs	0.96	1.00	0.98	258
Home Entertainment	1.00	0.50	0.67	4
Home Furnishing	0.96	1.00	0.98	179
Home Improvement	0.94	0.85	0.89	20
Jewellery	1.00	1.00	1.00	875
Kitchen & Dining	0.96	0.96	0.96	141
Mobiles & Accessories	0.98	0.99	0.98	287
Pens & Stationery	0.90	0.72	0.80	74
Pet Supplies	1.00	0.75	0.86	4
Sports & Fitness	0.97	0.83	0.89	35
Sunglasses	1.00	1.00	1.00	11
Tools & Hardware	0.94	0.94	0.94	104
Toys & School Supplies	0.82	0.90	0.86	71
Watches	1.00	0.98	0.99	118
eBooks	1.00	0.33	0.50	3
accuracy			0.98	4914
macro avg	0.96	0.88	0.91	4914
weighted avg	0.98	0.98	0.98	4914

Fig. 11. Classification report for TFIDF vectorizer

The below Fig.12 shows the QR generation.

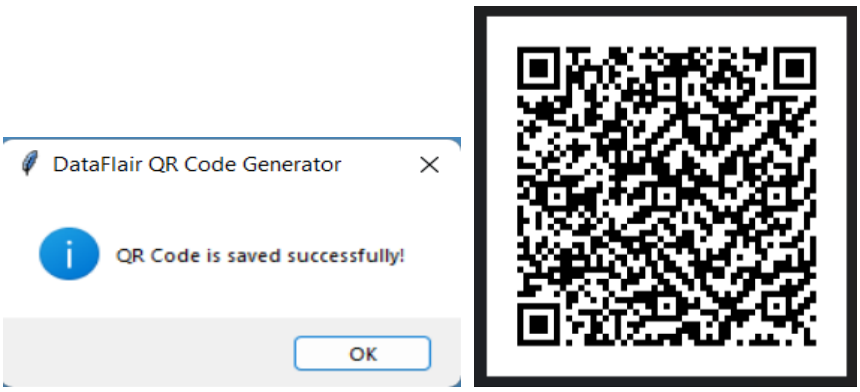

Fig. 12. Successful QR generation and QR

The home page shown in below Fig.13 consists of header which is having the store and about links. This is created using division tag with header and footer sections. Html, CSS and JS are used here and Bootstrap utilities are also used here. This webpage is screen fit for any type of resolution the user is using and can be simply scanned.

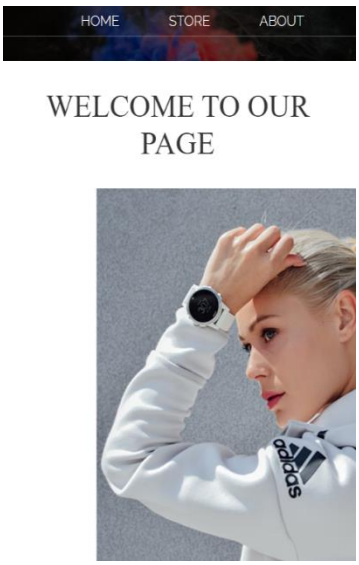

Fig. 13. Homepage of website

The below Fig.14 shows the mobile view of website and Fig.15 shows the feedback rating page,

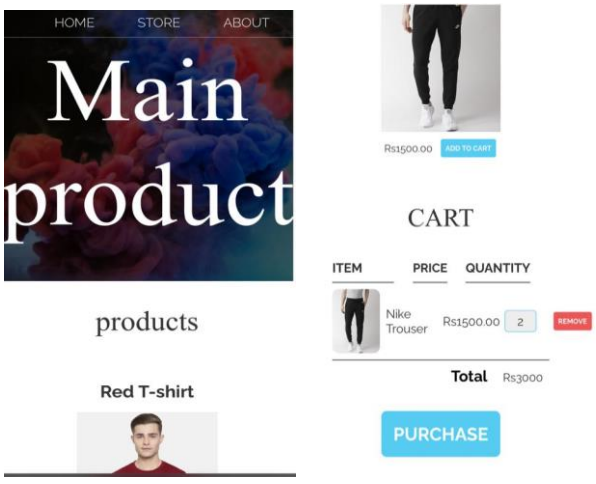

Fig. 14. Mobile view of website

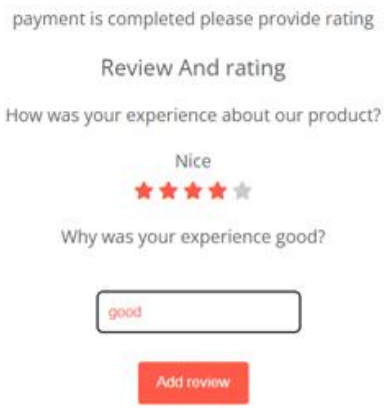

Fig. 15. Feedback page after purchase

Sentimental analysis for rating shown in below Fig.16:

	product title	Customer name	Rating	comment	Category	typeofreview
0	red tshirt	Rishikumar Thakur	5	product is very good	tshirt	1.0
1	red tshirt	Raza ji	4	fits good	tshirt	1.0
2	hrx shoes	Vaibhav Patel	4	material is excellent	shoes	1.0
3	hrx trouser	ishanth rajput	4	fantastic	trouser	1.0
4	hrx shoes	sundar tehar	5	good fit	shoes	1.0
5	hrx trouser	Arun Yajurvedi	3	good quality of product	trouser	1.0
6	hrx shoes	Irukulla bharath	3	good fit	shoes	1.0
7	hrx socks	rithwik	4	nice quality	socks	1.0
8	hrx trouser	siraj	5	worthy	trouser	1.0
9	hrx trouser	Sharad	5	good	trouser	1.0
10	hrx socks	Rishikumar Thakur	4	elastic	socks	1.0
11	hrx socks	Raza ji	3	good	socks	1.0
12	hrx socks	Vaibhav Patel	5	better quality	socks	1.0
13	black watch	rahul	4	good looking	watch	1.0
14	red tshirt	hari	5	product is very good	tshirt	1.0
15	red tshirt	shankar	4	fits good	tshirt	1.0
16	hrx shoes	reddy	4	material is excellent	shoes	1.0
17	hrx trouser	hemanth	4	fantastic	trouser	1.0
18	hrx shoes	sunil	5	good fit	shoes	1.0
19	hrx trouser	shetty	1	bad	trouser	0.0
20	hrx shoes	vicky	3	good fit	shoes	1.0
21	hrx socks	eswar	2	too costly socks	socks	0.0

Fig. 16. positive or negative review based on comment

The below Fig.17 and Fig.18 represents the classification report using SVM and Na ïve Bayes

```
print(classification_report(Test_Y,predictions_SVM))
```

```
              precision    recall  f1-score   support

           0       0.00      0.00      0.00         1
           1       0.83      1.00      0.91         5

    accuracy                           0.83         6
   macro avg       0.42      0.50      0.45         6
weighted avg       0.69      0.83      0.76         6
```

Fig. 17. Classification report using SVM

```
print(classification_report(Test_Y,predictions_NB))
```

	precision	recall	f1-score	support
0	0.00	0.00	0.00	1
1	0.83	1.00	0.91	5
accuracy			0.83	6
macro avg	0.42	0.50	0.45	6
weighted avg	0.69	0.83	0.76	6

Fig. 18. Classification report using Naïve Bayes

The below Fig.19 and Fig.20 represents the variation of actual vs predicted by Naïve Bayes and SVM.

```
print(predictions_NB)
print(Test_Y)
```

```
[1 1 1 1 1 1]
[1 1 1 1 0 1]
```

Fig. 19. actual vs predicted by Naive bayes

```
print(predictions_SVM)
print(Test_Y)
```

```
[1 1 1 1 1 1]
[1 1 1 1 0 1]
```

Fig. 20. actual vs predicted by SVM

5. Conclusion

In this, generating a QR code to overcome the limitation of existing one. QR code is limited to specific individual product and is not efficient to access the data in a faster manner. New QR code is developed to make the flexibility to users to access all the products in an efficient manner and error free. Our software provides a clear understanding of how to access the products using QR codes, and the webpage can be changed using a PHP connection to make it more safe for the users. Most software-generated QR codes have been black since a few years ago. These days, vivid and distinctive QR codes are used in the majority of advertisements. With QR codes, there is no need to be concerned about localised damage, the QR code's section being rewritten, or deleted because scanning devices can still recognise their entire information record. Right now, various websites offer free colourful QR code services to users. Making a specific QR code for a restaurant is also simple.. We have only ever seen a single QR code for a single product up until now, but the singer will soon provide the whole purchasing experience with just a single scan. Additionally, it is produced using cost-free servers.

References

[1] Bala Krishna, M., & Dugar, A. (2016). Product authentication using QR codes: a mobile application to combat counterfeiting. Wireless Personal Communications, 90(1), 381-398. DOI:10.1007/s11277-016-3374-x

[2] Tkachenko, I., Puech, W., Destruel, C., Strauss, O., Gaudin, J. M., & Guichard, C. (2016). Two-level QR code for private message sharing and document authentication. IEEE Transactions on Information Forensics and Security, 11(3), 571-583. https://doi.org/10.1109/TIFS.2015.2506546

[3] Singh, S. (2016). QR code analysis. International Journal of Advanced Research in Computer Science and Software Engineering, 6(5).

[4] Xu, S. (2018). Bayesian Naïve Bayes classifiers to text classification. Journal of Information Science, 44(1), 48-59. https://doi.org/10.1177/0165551516677946

[5] Verma, C., Stoffová, V., & Illés, Z. (2018). An ensemble approach to identifying the student gender towards information and communication technology awareness in European schools using machine learning. Int. J. Eng. Technol, 7(4), 3392-3396. doi: 10.14419/ijet.v7i4.14045

[6] Bordoloi, D. J., & Tiwari, R. (2014). Optimum multi-fault classification of gears with the integration of evolutionary and SVM algorithms. Mechanism and Machine Theory, 73, 49-60. https://doi.org/10.1016/j.mechmachtheory.2013.10.006

[7] Beniz, D., & Espindola, A. (2016). Using Tkinter of python to create a graphical user interface (GUI) for scripts in LNLS. WEPOPRPO25, 9, 25-28.

[8] Parmar, P. S., Biju, P. K., Shankar, M., & Kadiresan, N. (2018, September). Multiclass text classification and analytics for improving customer support response through different classifiers. In 2018 International Conference on Advances in Computing, Communications, and Informatics (ICACCI) (pp. 538-542). IEEE. http://dx.doi.org/10.1109/ICACCI.2018.8554881

[9] Fang, X., & Zhan, J. (2015). Sentiment analysis using product review data. Journal of Big Data, 2(1), 1-14. DOI 10.1186/s40537-015-0015-2

[10] Yadav, P., & Barwal, P. N. (2014). Designing responsive websites using HTML and CSS. International journal of scientific & technology research, 3(11), 152-155.

[11] Baid, P., Gupta, A., & Chaplot, N. (2017). Sentiment analysis of movie reviews using machine learning techniques. International Journal of Computer Applications, 179(7), 45-49. doi:10.5120/ijca2017916005

[12] Chatterjee, N., Chakraborty, S., Decosta, A., & Nath, A. (2018). Real-time communication application based on android using Google firebase. Int. J. Adv. Res. Comput. Sci. Manag. Stud, 6(4).

[13] Chintada, K. R., Yalla, S. P., & Uriti, A. (2021, November). A Deep Belief Network Based Land Cover Classification. In 2021 Innovations in Power and Advanced Computing Technologies (i-PACT) (pp. 1-5). IEEE. https://doi.org/10.1109/i-PACT52855.2021.9696524

[14] Uriti, A., Yalla, S. P., & Chintada, K. R. (2021, November). An Approach of Understanding Customer Behavior with an Emphasis on Rides. In 2021 Innovations in Power and Advanced Computing Technologies (i-PACT) (pp. 1-5). IEEE. https://doi.org/10.1109/i-PACT52855.2021.9696837

[15] Yalla, S.P., Uriti, A., Sethy, A. (2022). GUI implementation of modified and secure image steganography using least significant bit substitution. International Journal of Safety and Security Engineering, Vol. 12, No. 5, pp. 639-643. https://doi.org/10.18280/ijsse.120513

Authors' Profiles

Uriti Archana is working as Assistant Professor in the Department of Information Technology at GMR Institute of Technology, Rajam. Education qualification: B.Tech from JNTUK, M.Tech from JNTUK. Currently Pursuing Ph.D in Machine Learning and Computer Vision. Research interests include Big Data Analytics, Networking and Machine Learning.

Yalla Surya Prakash is working as Assistant Professor in the Department of Information Technology at GMR Institute of Technology, Rajam. Education qualification: B.Tech from JNTUK, M.Tech from JNTUK. Currently Pursuing Ph.D in Cyber Security. Research interests include Big Data Analytics and Network Security.

How to cite this paper: Archana Uriti, Surya Prakash Yalla, "Exploration on Quick Response (QR) Code Behaviour in Commerce based Platforms Using Machine Learning", International Journal of Information Engineering and Electronic Business(IJIEEB), Vol.15, No.5, pp. 1-12, 2023. DOI:10.5815/ijieeb.2023.05.01