Webology (ISSN: 1735-188X)
Volume 18, Number 6, 2021

Analysis Of User Behavior In Social Media Using Business Process Re- Engineering And Learning Technique

Vivek Shukla[1] , Rohit Miri[2]

[1,2]Department of Computer Science and Engineering, Dr. C. V. Raman University, Bilaspur Chhattisgarh, India.

Abstract

Social networking is quite prevalent in today's globe. People are using social networking sites to purchase a variety of items. Social media facilitates the formation of links among individuals of various backgrounds, resulting in a strong social structure. A significant result of this structure is the generation of massive amounts of data, which offers customers an exceptional service value proposition. BPR (business process reengineering) is a technique that may help businesses enhance quality, customer service, save costs, and become market leaders.BPR has been acknowledged as one of the most essential strategies for improving organizational performance in all business process performance measurements. BPR aimed to help organizations fundamentally rethink how they do their work to improve customer service, cut operational costs, and become world-class competitorswith the help of Sentiment analysis (SA).SA is one of the data mining types that estimates the directionof personality's sentiment analysis within natural language processing. In the implementation and work, sentiment analysis using the Twitter dataset is performed to assess the accuracy of several classifiers such as RF, LR, MNB, and CNB.

Keywords: Business Process Reengineering, Social media, Information systems, data analytics, Behavior analysis

1. Introduction

Organizational structure and behavior must now be considered to aid adaptation and evolution in an environment that is dynamic and changing at an accelerated rate [1]. Organizational changes have come as a surprise recently, despite being predicted in the past. As modern technologies emerge corporate processes become more globalized, and consumer needs to evolvethe organization's position in the market that changes day by day [2]. The goal is to increase the performance of the companythat provides exceptional work while keeping costs to a minimum and providing added value to the client by fully knowing their needs. Therefore, in a world of innovative technology, modifications, and powerful rivals need to be effectively and continuously rebuilt to achieve calculated and effectiveaccomplishment. The inefficiency of corporate procedures and a lack of innovation are to blame for an organization's strategic failures with major ramifications for businesses and their competitiveness [3].

Business Process Re-engineering (BPR) is described as a basicchanging and fundamentallyrestructuring of business procedures to generate significant enhancements in all implementation measures such as price, speed, value, and maintenance [4]. Every corporate and public entity is also required to employ BPR or is seeking other techniques that provide the same goals. Even though many organizations accepted the notion of BPR initiatives, only a handful of them succeeded, but the others failed through a superiorcollapse rate (e.g., 70%) [5].Various elements influence the performance of the BPR, which would be discussed in depth beneath, and one of these aspects is an awareness of the environment in which the business development operates. Therefore, businesses need knowledge management approaches and the addition of learning management standardsto comprehend the environment, which containsmethods, people, employees, customers, and implements [6]. The ontology encompasses a variety of ideas and categories, and it was created to better the understanding of the association's structure and the link among the

company's objectives it is also utilized in the learningfield [7]. A knowledge map is a description of information that illustrates links between knowledge sources by leveraging the comparison of maps to depict a specific location. It is a learning management method that is utilized for a variety of determinations, including locating causes of information or occasions for dataformation, increasing their contribution and how they interrelateinside the association, identifying skills, and determining the terms of reference [8].

The organization's task consists of creating massive amounts of data that provide clients with a one-of-a-kind value scheme. With so much information available, customers may find it difficult to get useful information when they need it. People use social networking sites for a number of purposes, including interactions, evaluating, and proposing products, office management, and getting up to date on current events. When most people think of social media, they see sites and apps such as WhatsApp, Facebook, YouTube, Twitter, Pinterest, LinkedIn, and Instagram. These are just a few social media examples. These applications depend on user-generated content and are important in a number of situations, including purchasing and selling habits, commercial business, governmental problems, scheme entrepreneurship, and so on [9].Facebook has the enviable importance of being the market administrator in the social media globeincluding 1.97 billion monthly clients as of April 2017 [10][11].

1.2 Business Process Re-engineering (BPR)

BPR has produced various benefits, including lower costs and higher productivity, better products, and more customer satisfaction as a result of its widespread usage since 1990. Each definition of process re-engineering places a different focus on the same basic idea: streamlining [11]. To achieve significant cost, quality, and service gains, BPR entails a complete rethink of procedures. Re-engineering is the process of revitalizing a company development that begins with ideas but does not result in anything being accepted.Supply chain traceability information flows may be managed using the BPR approach, which is a computer-based system [12]. By automating or replacing an existing legacy system, the implementation of an information system can be seen as a business process reengineering strategy as shown in Figure 1[13].

Figure 1:Flowchart of business process re-engineering [13]

1.3 Sentiments Analysis

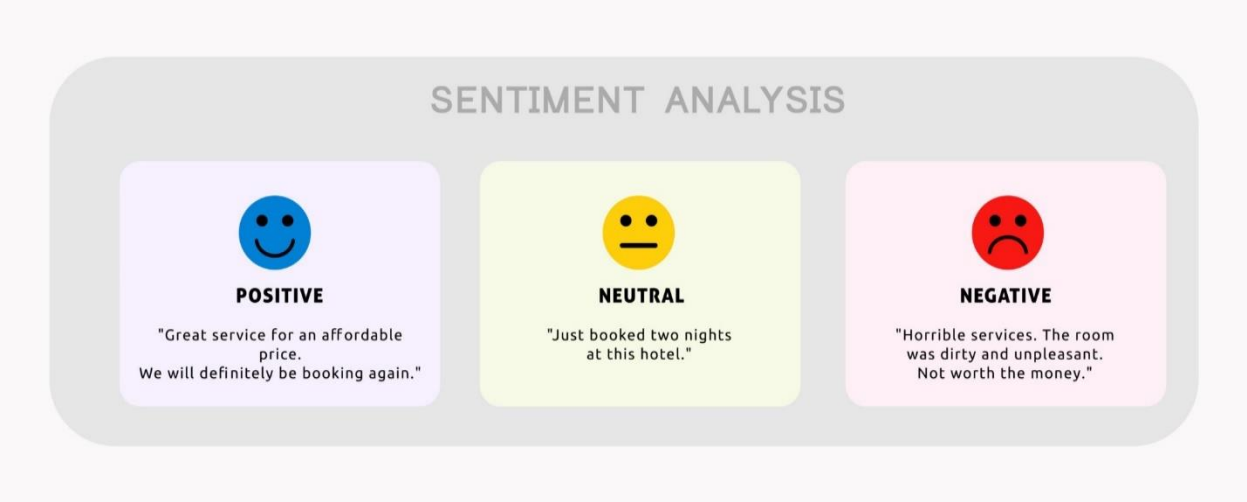
Sentiment analysis (SA) is the method of managing feelings, views, and subjective text [13]. SA provides understanding information on public opinions by analyzing various tweets and reviews. It is a proven technique for predicting numerous substantial events, for example, movie box office success and general elections [14]. Public reviews are applied to assess a specific item, such as a human, product, or place, and maybe discovered on many

websites such as Amazon and Yelp. Opinions may be classified as negative, favorable, or neutral. The goal of SA is to identify the expressive objective of client evaluations automatically [15]. The increased need for evaluating and organizing hidden data by social media in the form of unstructured data has increased the demand for sentiment analysis [16]. As it analyses various evaluations and tweets, SA provides comprehensible information related to popular opinions. It is a proven effective technique for forecasting many critical events such as general elections and box office movies [17].

Twitter, SMS, and other social networks have piqued the attention of the research business and community as vital sources of real-time opinion [18]. SA aggregates views as a) positive, b) neutral, or c) negative expression of the opinion holder [19]. Every day, a million tweets are sent out on a variety of topics. Two major difficulties in analyzing tweets are linguistic flexibility in expressing and topical variety in content. Various twitter sentiment analyzers depend on numerous sentiment lexicons to provide features to classify models or to calculate sentiment scores.

Figure 2 describe sentiment analysis diagram.

1.4 Features of Sentimental Analysis



Sentiments uses polarity and combinations to create a range of featured values for example trigrams and bigrams. As a result, feelings are evaluated as both negative and positive characteristics utilizing a variety of support vector machines and training methods. In sentiment analysis, neural networks are used to calculate label belongingness. The conditional relationships among multiple edges and nodes of an acyclic graph controlled by Bayesian networks are utilized to aid data extraction at the context level. Learning and data accuracy may be accomplished on social media platforms by optimizing words and phrases. Data tokenization is applied to generate negative and positive features of data at the word root level. Techniques are being developed to reduce SA mistakes to accomplish a better degree of accuracy in social media data [20].The 'Hidden Markov model' [20] is used for the optimum SA of paragraphs and phrases. The optimization of words and phrases leads to quicker learning, which improves social media data accuracy. Tokenization of data at the word root level aids in the creation of positive and negative aspects of information.

This paper is distributed into 4 sections, the 1st section contains the brief introduction of the paper and 2nd section contain the related work of various authors. The 3rd section contains the research methodology and implementation results, and the final section contains the conclusion.

2. Review of literature

The following study discusses in depth the preceding Analysis of user behavior in social media using business process re- engineering and learning technique. Various researchers explain their findings, as indicated below.

Dwipriyokoet al., (2020)[21]stated that every company's Enterprise Architecture implementation transitions include a BPR phase. One of the recommended practices in BPR is to limit the transition period to no more than 5 years. Enterprises in transition will see progressive changes in BPR, starting with partial unit changes and progressing to complete structural and functional changes. However, one successful New Generation Cooperatives Enterprise has been in the Partial BPR phase for more than 5 years.

Bitkowska et al., (2020) [22] stated that Integration of BPM and Knowledge Management (KM) concepts is a tough research issue these days, but it should be investigated in contemporary enterprises simultaneously. The issue for modern process organizations is the continual collecting of data and its professional use in order to gain a realistic advantage and maintain a balanced marketplace. Despite the growing interest among academics and practitioners, this topic has few publications. The major purpose is to determine the relationship between BPM and Knowledge Management.

Kapooret al., (2018) [23] has been developed as a consequence of communication platforms that allow the development of interactionsamongclients from various environments. Analysis and verdict-creating are supported by User-Generated Content (UGC).As a result, several studies have looked at the sustainability of virtualgroups and social media as anadvertising tool, butanotherhasconcentrated on the hazards and gains of utilizing social media in the workplace. Prior studies have looked on the usage of social media for communicationdistribution and for seeking and/or helping during crucial events. Public administration and Politics as well as the contrast of conventional and social media are examples of other situations. This study's completeassessment of the prevailinginformation that mayhelpthe upcomingscholarsstated terminationbut providing lucrative research avenues to help define this new field's research.

AbdEllatif et al., (2018) [24]stated that BPR has been highlighted as one of the highly essential strategies for organizational developments in all industry process execution measurements. However, significant failure rates of 70% have been observed while utilizing it. The extremely crucial cause for the collapse is the concentration on the procedure itself, independent of the adjacentecosystem or the company's understanding. Other issues include anabsence of implements to identify the root effects of discrepancies and incompetence.

Tadesse et al., (2018) [25]stated that with the rise of social networks, a plethora of techniques to defining individuals' personalities based on their social activities and language use patterns have emerged. Various machine learning methods, data sources, and feature sets need different techniques. The purpose of this study is to look at the predictability of Facebook users' personality characteristics using multiple features and metrics from the Big 5 model. Using the myPersonality project data set, we investigate the prevalence of social network structures and linguistic elements in relation to personality interactions. We evaluate and contrast four machine learning models, as well as examine the relationship between each feature set and personality attributes. Even when evaluated on the identical data set, the personality prediction system based on the XGBoost classifier surpasses the average baseline for all feature sets, with the greatest prediction accuracy of 74.2 percent. For the extraversion characteristic, the best prediction performance was attained by applying the individual social network analysis features set, which had a better personality prediction accuracy of 78.6 percent.

Bhaskar et al., (2018)[26]stated thatfor radical redesign and enhancement of business processes, BPR has become a prominent change management strategy. It encourages businesses to accomplish things more efficiently in order to improve overall quality. Unfortunately, more than 70% of BPR deployments are considered to have failed to produce projected benefits owing to a lack of appropriate framework and methodology. As a result, the author attempted to evaluate several BPR frameworks and methodologies in order to build an appropriate framework and methodology as well as to fill a vacuum in the literature. The study's secondary objectives are to investigate the BPR variables that may impact performance and to investigate the amount of BPR adoption in Indian manufacturing firms. Based on the design of contemporary BPR frameworks and techniques recognized in the literature, this research proposes a common BPR framework and methodology for organizations.

Eshuiset al., (2016)[27]stated that to defineflexible business processes, data-centric BP models combine the data and control flow. However, since the global process is frequently spread across many data components and potentially described in a declarative manner, it may be tricky to anticipate the genuine behavior of a data-centric model. As a result, envisage a "data-centric process modeling" method in which the process's default behavior is first defined in a traditional, imperative process notation, then converted into a declarativedata-centric process model, which can be more developed into a full model.

Ferraraet al., (2014) [28]analyzed that web data extraction is a significant subject that has been investigated using many scientific approaches and in a wide variety of applications. Various ways to data extraction from the Web have been developed to address challenges and work in adhoc contexts. Otherssignificantly rely on methods and algorithms established around information extraction. At Web Data Extraction methods and the Social Web level enable the collection of massive amounts of structured data created and distributed by online social network users, social media, and web 2.0, providing unparalleled opportunity to examine human behavior on a massive scale. It is also exploring the prospect of cross-insemination, i.e., the chance of reusing web data extraction methods that were initially developed to function in a certain domain in other domains.

Bhaskar et al., (2014) [29]explained that BPR is a method that may improve firms enhance value, client benefit, decrease controlling expenses, and develop industry directors. BPR might be a significant strategic instrument for sustaining useful improvement for foreign or Indian business organizations (private and public). The intent of this study was to gather and evaluate previous work in the subject of quality managing and BPR. The primary goal of this research was to provide a complete review of the whole evolution of the BPR idea, theories, methodologies, problems, and results. Based on prior research investigations, it was established that using BPR in the Indian industrial segment is not complicated, yet although conditions change and the causes for variation change in this segment. There is even an essential for a private and globally recognized model for BPR, as well as a technique that is widely applicable.

Dinçkanet al., (2013) [30] examined thatglobalization and rising competition force most firms to become more creative and employ change-based methods. BPR is one of the best common transformationmanaging methodologies that has lately received a lot of attention in the world of change. Even however there have been both successful and failed cases reported in the fiction, BPR has been hailed by business circles as a critical management tool for achieving extraordinaryenhancements and governmentaleffectiveness if executed properly and carefully. In this research, business process reengineering is targeted at producing remarkable improvements on the existing business processes of a corporation in the private sector by employing simulation technology. In this case, literature study was conducted initially to fully comprehend the issue. Then, several BPR approaches from the literature were explored, as well as enablers of BPR that make change possible and function as vehicles for process transformation.Table 1 illustrates the summary of the above work.

Table 1:Summary of Related Work

S.no.	Author	Year	Outcomes
1.	Dwipriyoko et al., [21]	2020	New Generation Cooperatives Enterprise has been in the Partial BPR phase.
2.	Bitkowska et al., [22]	2020	The issue for modern process organizations is the continual collecting of data and its professional use in order to gain a realistic advantage and maintain a balanced marketplace.

3.	Kapoor et al., [23]	2018	Prevailing information that may help the upcoming scholars stated termination but providing lucrative research avenues to help define this new field's research.
4.	AbdEllatif et al.,[24]	2018	Significant failure rates of 70% have been observed while utilizing it.
5.	Tadesse et al., [25]	2018	XGBoost classifier surpasses the average baseline for all feature sets, with the greatest prediction accuracy of 74.2 percent.
6.	Bhaskar et al., [26]	2018	More than 70% of BPR deployments are considered to have failed to produce projected benefits owing to a lack of appropriate framework and methodology.
7.	Eshuis et al.,[27]	2016	Data-centric process modeling" method in which the process's default behavior is first defined in a traditional, imperative process notation.
8.	Ferrara et al., [28]	2014	The chance of reusing web data extraction methods that were initially developed to function in a certain domain in other domains.
9.	Bhaskar et al., [29]	2014	BPR might be a significant strategic instrument for sustaining useful improvement for foreign or Indian business organizations.
10.	Dinçkan et al., [30]	2013	BPR has been hailed by business circles as a critical management tool for achieving extraordinary enhancements and governmental effectiveness.

3. Background study

Li et al. (2021) [31] developed a latest approach on the classification and extraction of the online dating service comments. Using social cognition theory, the author aimed to understand and obtain the emotion idea of each emotional evaluation, as opposed to standard emotional assessment, which mostly concentrates on product designation.The author took a sample of 4,300 comments from dating websites with significantly negative/positive emotions and analyzed them using three machine learning techniques.The author uses numerous machine learning techniques,dictionary-based sentiment analysis and sentiment analysis while analyzing and evaluating the proficiency of user behavior research.The author discovered that combining machine learning with a lexicon-based strategy can reach greater precision than any other sort of emotionassessment.

4. Problem Formulation

Most of the organization is participating in BPR, whether they are aware of it or not. Also, the goal of enterprises to overcome competitive gaps and achieve higher performance standards is a major factor in the adoption of the BPR approach. Current study aims to improve BPR by using user/employee behavior analysts gathered from social media networks. As a part of the data center BPR model, data is used for process identification and mapping, as well as to verify the change made in each process. Classification of user and employee behavior is based on benefit risks, trusts,

and distrusts, and PSI (Positive social influence)/NSI (Negative social influence) in the whole BPR system architecture, with a variety of machine learning classifiers used to provide the best categorized results.

5. Research methodology

This work is based on designing of BPR in order to increase the efficiency of the employees in the organization. The BPR model is based on the classification of the employ behavior that is the impact of positive and negative social influence.The current BPR methodology is distinct from conventional BPR approaches because of its flexibility and more integrated approach. Sentiment analysis model is used in the proposed work for analyzing the user.

- **Sentiment classification model**

As part of multi-emotivity learning technique, each of the three basic classifiers is based on one of the three categoriesof characteristics (benefit/risk, trust/distrust, PSI/NSI). Using a logistic regression model, the three classifiers are concatenated. logistic regression employs a logistic function to develop a binary dependent variable. When describing the data, this can be used to highlight the relationship between one or more nominal, ratio-level, ordinal, causative circumstances, as well as the causal factor itself [31].

i. **Benefit/risk classifier:**In the context of online social data, benefits advantages may lead to specific adoption, but perceived risks resulted in the lack of adoption [31].

ii. **Trust/distrust classifier:**In online social data context, individuals who have significant trust in the service would provide to individual adoption, although distrust could result in no-adoption [31].

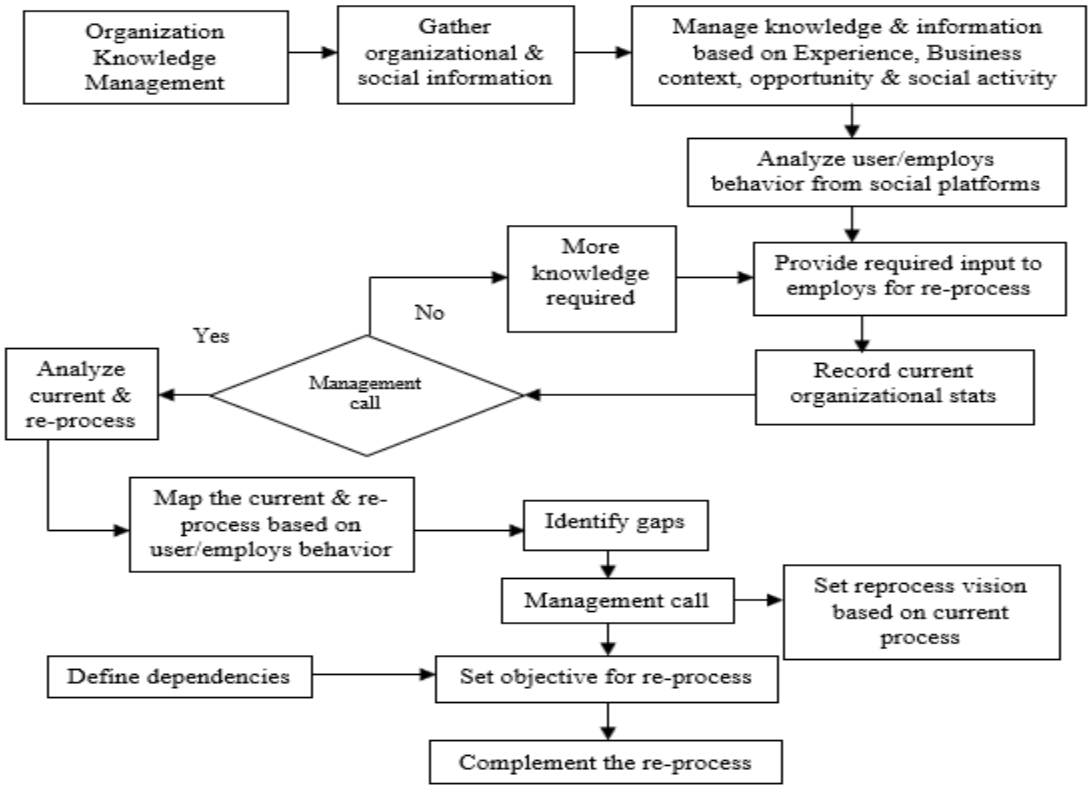
iii. **PSI/NSI classifier:** “Negative Social Influence (NSI)” is related with non-adoption while “Positive social influence (PSI)” may promote to specific adoption in the online environment [31].

6. Proposed methodology

This section of the paper contains implementation done using the proposed methodology:

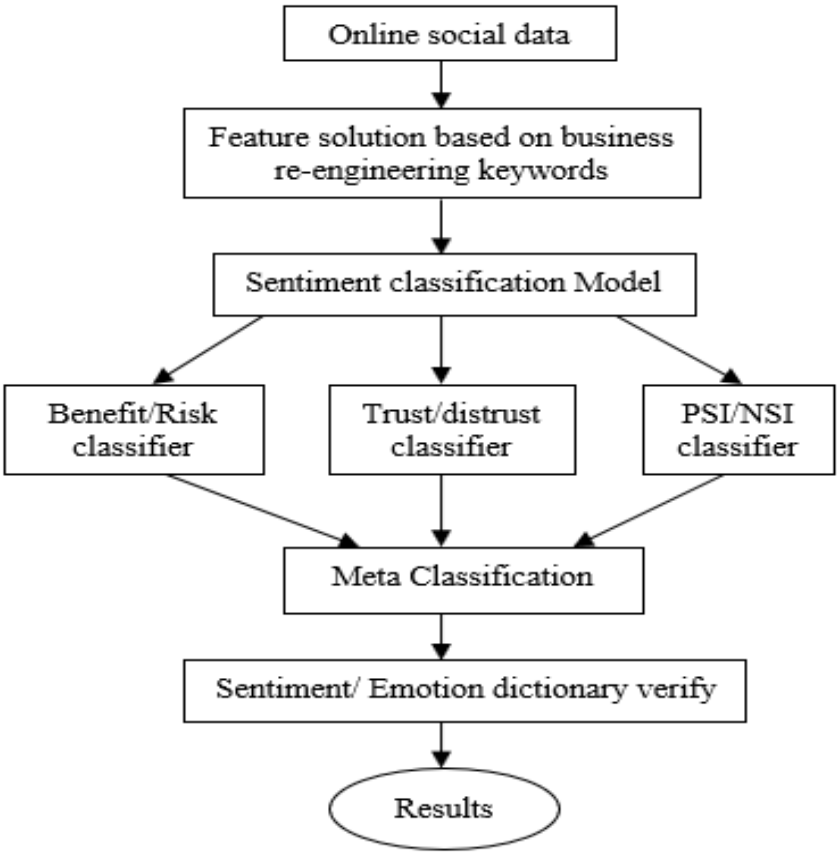
The proposed methodology is summarizing in two parts such as business process re-engineering and the user/employee behavior analysis.Figure 3 depicts the block diagram of the proposed methodology

Figure 3:Block diagram of the proposed methodology



The steps of the business re-engineering process is explaining below:

- **Step 1:**Firstly, a proper record of the organization is managed in a knowledge base.
- **Step 2:**Gather all the information of the organization that is required to analyze the behavior of the employs.
- **Step 3:**Manage all the information and the knowledge base on the basis of experience, business context, opportunity, and the social activity.
- **Step 4:**On the basis of the knowledge base, analyze the behavior of the employs with the help of different social media platforms.
- **Step 5:**After analyzing the behavior of the employee, provide the required input from the management side to the employs for re-processing. The in-detail process is depicted in Figure 4.



- The steps for the employs behavior analysis is summarizes below:
  a) The online social data is used in this research work. The dataset is created by gathering the information from the online social media sites.
  b) In this step apply the feature selection approach based on the business re-engineering keywords. The data that obtain from the previous step is in the unstructured form. Feature selection is used to convert the unstructured data in the structured form.
  c) In this step, sentiment classification model is created. The sentiment classification model is further divided in 3 different types such as trust/distrust classifier, Benefit/Risk classifier and PSI/NSI classifier.
  d) The output of the 3rd step is the input of the meta classification phase. All the outputs of the different classifier is combined in this step and convert it in a single output.
  e) At the end verify the emotions and the sentiments with the output of the meta classification and analyze the emotions that is positive for the organization or not.
- **Step 6:**Now, record the current statistics of the organization. On the basis of these records the higher management analyze the data of each employee inside the organization and decide if more knowledge is required or not.
- **Step 7:**If more knowledge is required, then again, repeat the process from the 5th step. If not, then analyze the current stats with the re-process.
- **Step 8:**Map the current and re-process data on the basis of the behavior of employees. After mapping, identify the gap about in which section the organization is lacking.
- **Step 9:**After the identification of gaps, the management discuss on the problems. On the basis of discussion, the management set new objective else set the reprocess vision based on the current process.

- **Step 10:**The objectives are set on the bases of the dependencies and complete the procedure of the business re-engineering.

7. Implementation tool and Results

This section of the paper contains implementation done using the proposed methodology:

7.1. Python Language

It is an actively semantic, object-oriented high-level, construed language. Its built-in high-level data structures, merged by dynamic linking and dynamic typing, produce this ideal for quick Implementation Development as a scripting language for connecting current elements. Python's simple-to-learn structure stresses readability that lowers software maintenance costs. Packages and modules are endorsed by Python that enables code reuse and software modularity [32].

7.2. Classifiers used in the implementation

a) Logistic Regression

The Logistic Regression (LR) is the popular linear classification algorithm. Logistic regression allows a relationship between an independent variable and dependent variables to form a multivariate regression. LR, which is the model of multivariate analysis, which is useful to predict the existence or in the absence of a function or consequence based on the values of a series of various predictor variables. The logistic regression model is applied to represent the probability characterizing the result of an experiment [33]. Maximum Entropy is another name for this approach. The logistic regression simply divides the data into two parts with the help of a line of regression as shown in Figure 5.

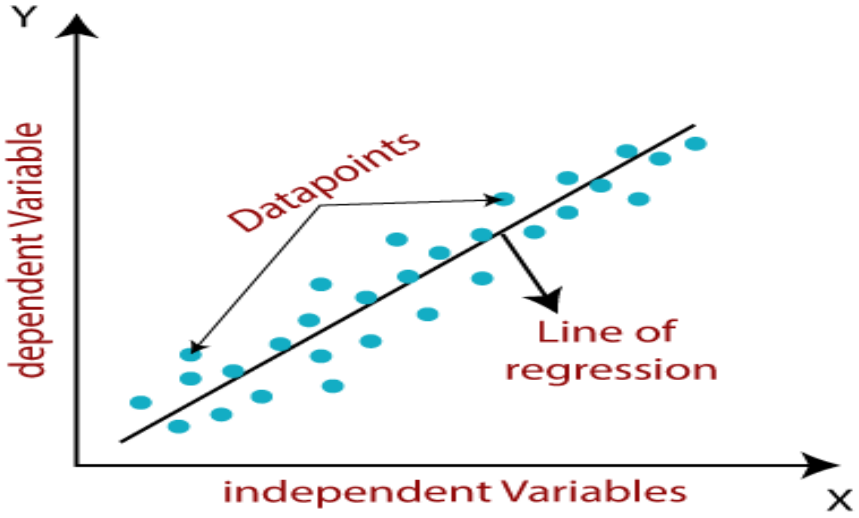


Figure 5: Logistic Regression [33]

b) Random Forest

Random forest is one of the most popular machine learning techniques that can be applied for regression and classification problems in machine learning algorithms. It is a collection of decision tree algorithms that may be applied for classification as well as regression. In this strategy, more trees often correlate to improve efficiency and performance. Extract a sample set of data points from a given training set using the bootstrap method. Create a decision tree based on the previous phase's findings and obtain the number of trees if the previous two phases are followed. Every tree that is planted would vote for a data point. Compute the majority voting of the decision tree classifiers [34]. The random forest's structure is seen in Figure 6.

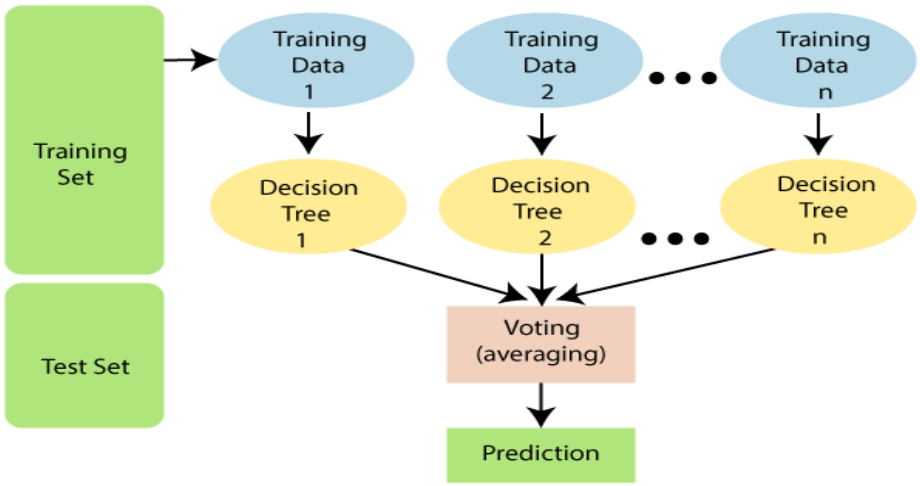


Figure 6:Random Forest [34]

c) Multinomial Naive Bayes (MNB)

The MNB algorithm is a common Bayesian learning method in Natural Language Processing (NLP). Using the Bayes theorem, the software estimates the tag of a text, such as an email or a newspaper piece. It computes the probability of each tag for a given sample and returns the tag with the highest chance. The MNB classifier is appropriate for classifying discrete features (e.g., word counts for text classification). Normally, integer feature counts are required for the multinomial distribution [35].

$$\hat{\theta}_{yi} = \frac{N_{yi}+\propto}{N_y+\propto n} \qquad (1)$$

Here "$N_y$is the total number of features of the event y, $N_{yi}$is the count of each feature (summary number of repetitions of a word in all spam messages), n is the number of features and $\propto$ is the smoothing Laplace parameter to discard the influence of words absent in the vocabulary".

d) Complement Naive Bayes (CNB)

CNB is a variant of the classic Multinomial Naive Bayes method. MNB does not perform well on skewed datasets. Imbalanced datasets are those in which the number of instances of one class exceeds the number of examples of other classes. Instead of assessing the likelihood of an item belonging to a certain class, CNB calculates the chance of the item belonging to all classes. This is the literal definition of the term, complement, and so it is abbreviated as CNB [36].

$$\hat{\theta}_{ci} = \frac{N_{ci}+\propto_i}{N_c+\alpha} \qquad (2)$$

Here N_c is the total number of words in the opposite class, N_{ci} is the repetitions of a word in the opposite class.

Result 1:

Figure 7 shows the input dataset which containstextId-unique ID for each piece of text, text-the text of the tweet, selected text-the text that supports the tweet's sentiment, and sentiment-the general sentiment of the tweet.This is done to extract the part of the tweet which provides the overall sentiment to the tweet.

Result 2:

	textID	text	selected_text	sentiment
0	cb774db0d1	I`d have responded, if I were going	I`d have responded, if I were going	0.0
1	549e992a42	Sooo SAD I will miss you here in San Diego!!!	Sooo SAD	2.0
2	088c60f138	my boss is bullying me...	bullying me	2.0
3	9642c003ef	what interview! leave me alone	leave me alone	2.0
4	358bd9e861	Sons of ****, why couldn`t they put them on t...	Sons of ****,	2.0

Figure 8 shows the twitter sentiment analysis in which blue colour shows the neutral sentiments, yellow colour shows

Figure 7: Input Dataset

the positive sentiment, and red colour shows the negative sentiment.

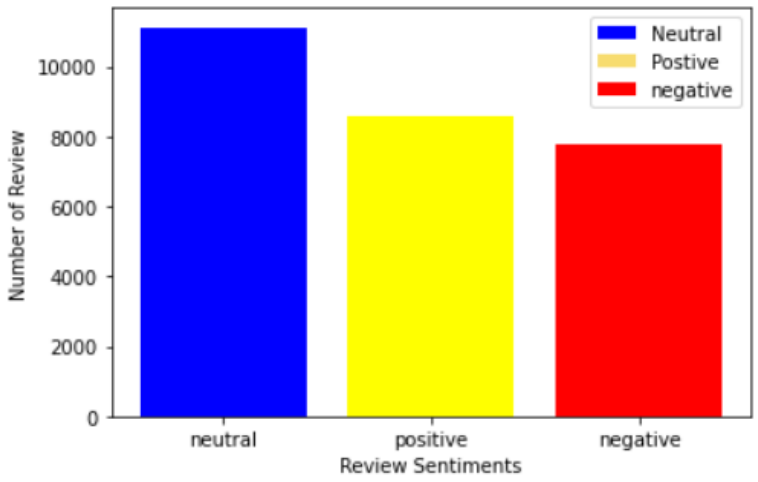


Figure 8: Twitter sentiment analysis

Result 3:

Figure 9 shows the labelling of sentiment in the dataset in which neutral sentiment denoted by "0", positive sentiment denoted by "1", and negative sentiment denoted by "2".

Result 4:

```
{0: 'NEUTRAL', 1: 'POSITIVE', 2: 'NEGATIVE'}
```

Figure 9: Labelling of sentiment

Figure 10 shows the preprocessing of data which is used to extract the feature from the selected text in which each selected text is denoted by the sentiment number.

	selected_text	sentiment
0	I`d have responded, if I were going	0
1	Sooo SAD	2
2	bullying me	2
3	leave me alone	2
4	Sons of ****,	2

Result 5:

Figure 10: Preprocessing of data

Figure 11 shows the logistic regression confusion matrix and its performance parameters which is calculated by the following method:

Precision (P): The proportion of relevant occurrences among the recovered examples is known as precision. The following equation is applied to determine precision:

$$P = \frac{TP}{(TP+FP)} \quad (3)$$

Where TP True Positive and FP stands for False Positive.

Recall (R): The proportion of relevant occurrences that were recovered is known as recall (also known as sensitivity). As a result, the relevance lies at the heart of both accuracy and memory. The following equation is used to calculate recall:

$$R = \frac{TP}{(TP+FN)} \quad (4)$$

Here FN stands for False-negative.

F1-score: Weighted average of recall and precision is called f-score.More important parameter than accuracy when having an uneven class distribution in data. It is calculated as follows:

$$\text{F1 score} = \frac{2*\text{Precision}*\text{Recall}}{\text{Precision}+\text{Recall}} \quad (5)$$

Accuracy: This is the ratio of true positives plus true negative to the true positives plus true negatives plus false positive plus false negative as shown in equation 4.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} (6)$$

```
Accuracy is :  90.07542200407039
Precision is :  94.55431355689309
Recall is :  97.89317507418399
fscore is :  96.19478058025952
array([[3299,  190,  290],
       [  71, 2267,  164],
       [ 404,   17, 1958]])
```

Figure 11: Logistic regression

Result 6:

Figure 12 shows the logistic regression error values according to the analysis mean absolute error is 0.25166, mean squared error is 0.42001, and root mean squared error is 0.64808.

```
Mean Absolute Error of LR: 0.25166767738023044

Mean Squared Errorof LR : 0.4200121285627653

Root Mean Squared Error of LR: 0.648083427162557
```

Figure 12: Logistic regression error values

Result 7:

Figure 13 shows the MNB confusion matrixin this MNB calculates each tag's likelihood for a given sample and outputs the tag with the greatest chance.

Result 8:

```
CM of MultinomialNB:
 [[3201  104   59]
 [ 552 1903   47]
 [ 862  194 1323]]
```

Figure 13: MNB confusion matrix

Figure 14 shows the performance parameters of MNB according to the analysis accuracy of MNB is 77.95, precision of MNB is 96.85, recall of MNB is 85.91, and f1 score of MNB is 90.70.

Result 9:

Figure 14: Performance parameters of MNB

```
Accuracy  of Multinomial NB is :  77.95027289266223
Precision is :  96.85325264750378
Recall is :  85.29176658673062
fscore is :  90.7055823179371
```

Figure 15 shows the MNB error values according to the analysis mean absolute error is 0.33220, mean squared error is 0.55560, and root mean squared error is 0.74539.

Result 10:

Figure 16 shows the confusion Metrix of CNB.

Result 11:

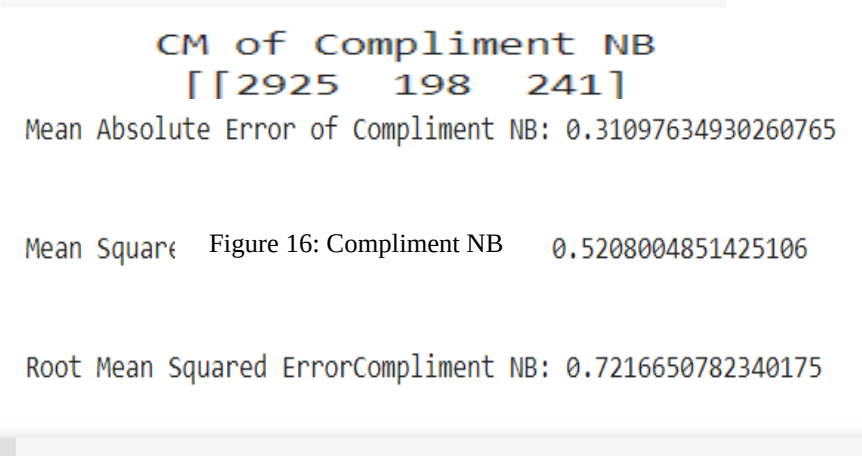


Figure 16: Compliment NB

Figure 17: Performance parameters of CNB

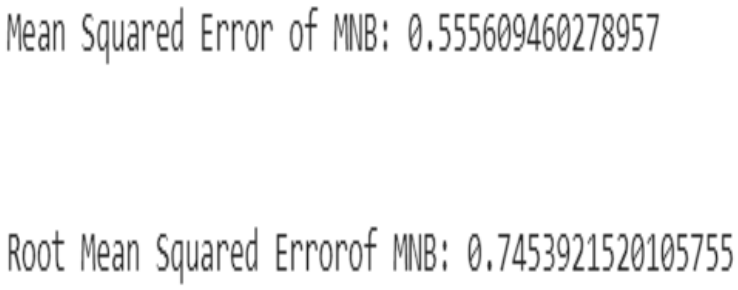


Figure 17 shows the performance parameters of CNB according to the analysis accuracy of CNB is 79.39, precision of CNB is 93.65, recall of CNB is 88.42, and f1 score of CNB is 90.65.

Result 12:

Figure 18 shows the CNB error values according to the analysis mean absolute error is 0.31097, mean squared error is 0.52080, and root mean squared error is 0.72166.

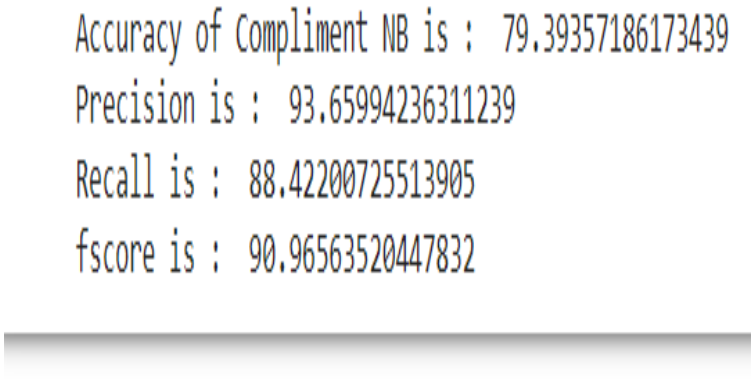


Figure 18: Error values of CNB

Result 13:

Figure 19 shows the confusion matrix of RF.

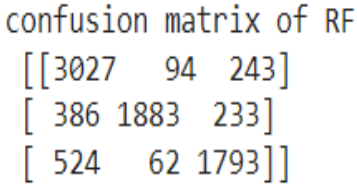


Figure 19: Confusion matrix of RF

Result 14:

Figure 20 shows the performance parameters of RF according to the analysis accuracy is 81.297, precision is 96.881, recall is 88.690, and f1 score is 92.65.

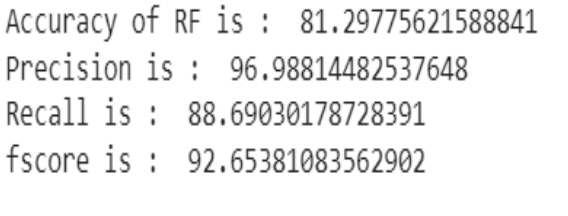


Figure 20: Performance parameters of RF

Result 15:

Figure 21 shows the RF error values according to the analysis mean absolute error is 0.28004, mean squared error is 0.46610, and root mean squared error is 0.628715.

```
Mean Absolute Error of RF: 0.28004851425106125

Mean Squared Error of RF: 0.4661006670709521

Root Mean Squared Error of RF: 0.682715656090405
```

Figure 21: Error value of RF

Result 16:

Bagging: Bagging, also known as bootstrap aggregation, is the ensemble learning method that is commonly used to reduce variance within a noisy dataset. In bagging, a random sample of data in a training set is selected with replacement—meaning that the individual data points can be chosen more than once. Figure 22 shows the accuracy of bagging.

```
Accuracy of Bagging is :  90.17779435534004
Precision is :  97.4715549936789
Recall is :  93.90986601705238
fscore is :  95.6575682382134
array([[3084,   80,  200],
       [ 200, 2207,  160],
       [ 350,   67, 1962]])
```

Result 17:

Figure 22: Bagging Accuracy

Booting: Boosting is an ensemble learning method that combines a set of weak learners into a strong learner to minimize training errors. In boosting, a random sample of data is selected, fitted with a model and then trained sequentiallythat is, each model tries to compensate for the weaknesses of its predecessor**.**Figure 23 shows the boosting accuracy.

```
Accuracy of Boosting is :  90.69247952345495
Precision is :  97.79110129378353
Recall is :  94.22316813621161
fscore is :  95.97398575410344
array([[3084,   80,  200],
       [ 200, 2207,  160],
       [ 350,   67, 1962]])
```

Figure 23: Boosting Accuracy

Result 18:

Voting:A Voting Classifier is a machine learning model that trains on an ensemble of numerous models and predicts an output (class) based on their highest probability of chosen class as the output. Figure 24 shows the voting accuracy.

```
voting accuracy
0.8999272197962155
```

Figure 24: Voting accuracy

Result 19:

Figure 25 shows the reviews on sentiments which contain negative, neutral, and positive sentiments.

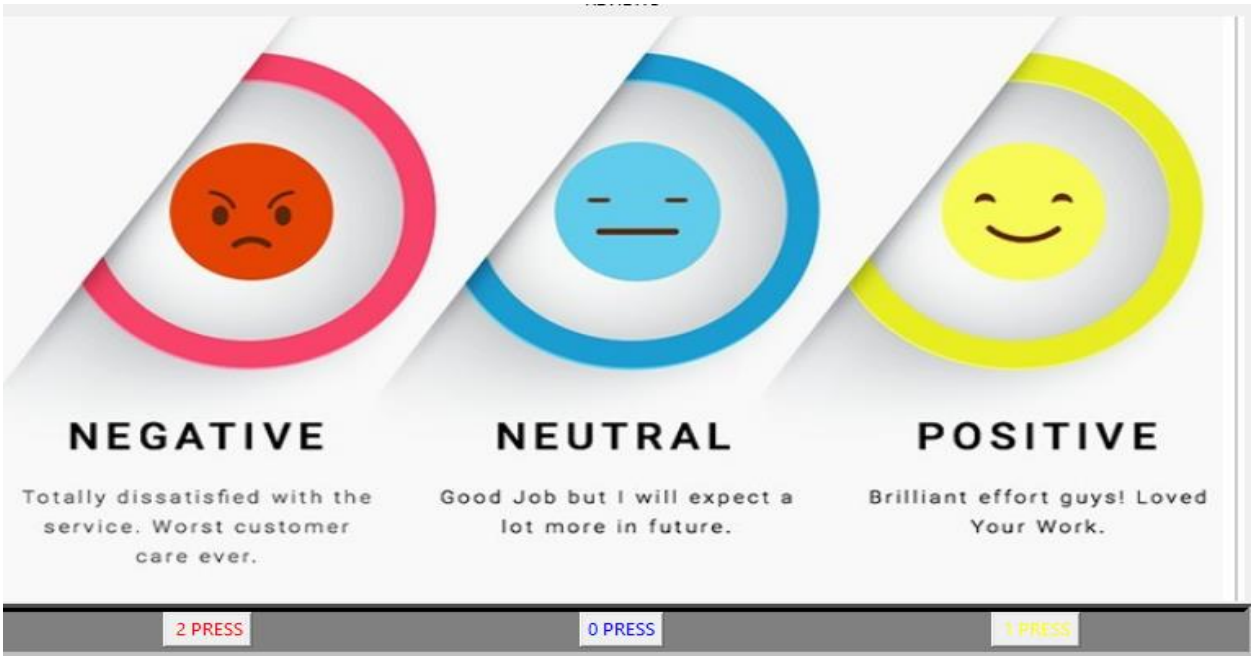


Figure 25: Reviews on sentiments

Result 20:

Figure 26 shows the further analysis which can be done for improving the quality of analysis.

```
Improve your quality
Enhance features
    1.Reduced cost
    2.No delay on dilvery time
    3.Should be Reliable
Satisfied with quality
```

Result 21:

Figure 27 shows the array which contains the output of the overall sentiment analysis based on parameters

```
predicted2

array([2, 0, 1, 2, 1, 0, 1, 2, 0, 1, 2, 0], dtype=int64)
```

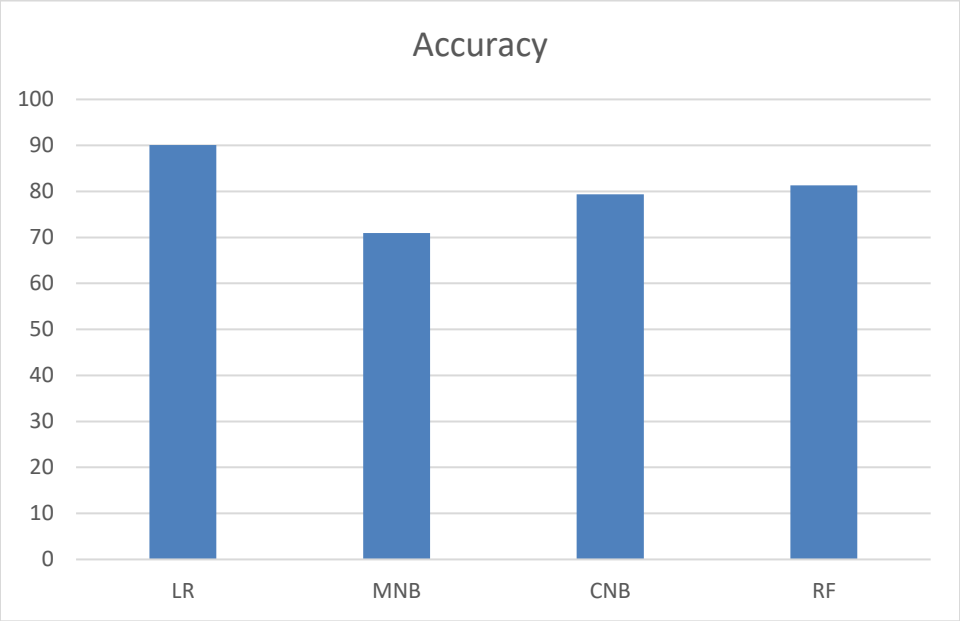
Figure 27: Output of the analysis

Result 22:

From the above analysis the Table 2 shows the classifier with their accuracy and Figure 28 shows the comparison graph of it according to which LR is having the higher accuracy which is 90.075.

Table 2: Classifier with their Accuracy

Classifier	Accuracy
LR	90.075
MNB	70.95
CNB	79.39
RF	81.29



Result 23:

Figure 28: Comparison graph

Table 3 below shows the comparison of the proposed work with the existing work. Here, in the proposed work logistic regression has enhanced accuracy as compared to random subspace in existing work [37]. The classifier accuracy with bagging and boosting algorithm accuracy is also more than the existing work.

Table 3: Comparison among existing vs proposed model

Parameter	Existing work [37]	Proposed Work
Technique Used	Random subspace	Logistic regression
Classifier Accuracy	89.2%	90.075%
Bagging Accuracy	82.5%	90.177%
Boosting Accuracy	85.2%	90.692%

8. Conclusion and Future Scope

Sentiment classification in social communication medium, notably online chat services, has sparked a lot of interest in the scientific community. Several BPR strategies have been created during the previous decade to improve reengineering undertakings and approaches. When dealing with complicated issues, the first BPR approaches were fairly simplistic and proved ineffectual at times. A data-driven process reengineering method was developed in this research to overcome these constraints and boost the success rate. In the implementation work sentiment analysis from twitter dataset is done in which accuracy of various classifier such as RF, LR, MNB, and CNB are determined and finally output based on sentiments is describe by the array. In future scope several models have appeared to improve the work of organizations' business processes through the use and development of business process re-engineering, but they have unexpected results, despite the fact that most organizations are using business re-engineering to solve all of their problems, develop their work properly and gain high quality. To examine the failure rates and explanations for any model, BPR requires statistical analytic techniques.

References

[1]. Tripathi, Rajesh Kumar, Subhash Chand Agrawal, and Anand Singh Jalal. "Real-time based human-fall detection from an indoor video surveillance." International Journal of Applied Pattern Recognition 5, no. 1 (2018): 72-86.

[2]. Messadi, Mahammed, Abdelhafid Bessaid, and Sihem Lazzouni. "New semi-automated segmentation approach of the left ventricle applied to cine MR images analysis." International Journal of Applied Pattern Recognition 5, no. 1 (2018): 40-54.

[3]. Rao, Lila, Gunjan Mansingh, and Kweku-Muata Osei-Bryson. "Building ontology-based knowledge maps to assist business process re-engineering." Decision Support Systems 52, no. 3 (2012): 577-589.

[4]. Singh, Pawan Kumar, Ram Sarkar, and Mita Nasipuri. "A comprehensive survey on Bangla handwritten numeral recognition." International Journal of Applied Pattern Recognition 5, no. 1 (2018): 55-71.

[5]. Ozcelik, Yasin. "Do business process reengineering projects payoff? Evidence from the United States." International Journal of Project Management 28, no. 1 (2010): 7-13.

[6]. Kavitha, N., and D. N. Chandrappa. "Vehicle tracking and speed estimation using view-independent traffic cameras." International Journal of Applied Pattern Recognition 6, no. 2 (2020): 163-176.

[7]. Suresh, P., and K. R. Radhika. "Nature inspired hybrid algorithms for binding shared key with user trait." International Journal of Applied Pattern Recognition 6, no. 3 (2021): 217-231.

[8]. Yoo, Keedong, Euiho Suh, and Kyoung-Yun Kim. "Knowledge flow-based business process redesign: applying a knowledge map to redesign a business process." Journal of knowledge management (2007).

[9]. Greenwood, Brad N., and Anand Gopal. "Research note—Tigerblood: Newspapers, blogs, and the founding of information technology firms." Information Systems Research 26, no. 4 (2015): 812-828.
[10]. Jayaram, M. A., Gayitri Jayatheertha, and Ritu Rajpurohit. "Time Series Predictive Models for Social Networking Media Usage Data: The Pragmatics and Projections." Asian Journal of Research in Computer Science (2020): 37-54.
[11]. Lister, M. "Essential Social Media Marketing Statistics For 2017," Wordstream." (2017).
[12]. Hammer, Michael, and James Champy. Reengineering the Corporation: Manifesto for Business Revolution, A. Zondervan, 2009.
[13]. Jin, Zhou, Yujiu Yang, Xianyu Bao, and Biqing Huang. "Combining user-based and global lexicon features for sentiment analysis in twitter." In 2016 International Joint Conference on Neural Networks (IJCNN), pp. 4525-4532. IEEE, 2016.
[14]. Maram, Balajee, J. M. Gnanasekar, Gunasekaran Manogaran, and Muthu Balaanand. "Intelligent security algorithm for UNICODE data privacy and security in IOT." Service Oriented Computing and Applications 13, no. 1 (2019): 3-15.
[15]. Anand, M. Bala, N. Karthikeyan, S. Karthick, and C. B. Sivaparthipan. "Demonetization: a visual exploration and pattern identification of people opinion on tweets." In 2018 International Conference on Soft-Computing and Network Security (ICSNS), pp. 1-7. IEEE, 2018.
[16]. Pandarachalil, Rafeeque, Selvaraju Sendhilkumar, and G. S. Mahalakshmi. "Twitter sentiment analysis for large-scale data: an unsupervised approach." Cognitive computation 7, no. 2 (2015): 254-262.
[17]. Singh, Jaspreet, Gurvinder Singh, and Rajinder Singh. "A review of sentiment analysis techniques for opinionated web text." CSI transactions on ICT 4, no. 2-4 (2016): 241-247.
[18]. Bharti, Santosh Kumar, Bakhtyar Vachha, R. K. Pradhan, Korra Sathya Babu, and Sanjay Kumar Jena. "Sarcastic sentiment detection in tweets streamed in real time: a big data approach." Digital Communications and Networks 2, no. 3 (2016): 108-121.
[19]. Maheswari, K. "Improving accuracy of sentiment classification analysis in twitter data set using knn." Int. J. Res. Anal. Rev 5, no. 1 (2018): 422-425.
[20]. Sivaparthipan, C. B., N. Karthikeyan, and S. Karthik. "Designing statistical assessment healthcare information system for diabetics analysis using big data." Multimedia Tools and Applications 79, no. 13 (2020): 8431-8444.
[21]. Dwipriyoko, Estiyan. "Partial Business Process Re-engineering in New Generation Cooperatives Enterprise Architecture Implementation." In Journal of Physics: Conference Series, vol. 1477, no. 5, p. 052033. IOP Publishing, 2020.
[22]. Bitkowska, Agnieszka. "The relationship between Business Process Management and Knowledge Management-selected aspects from a study of companies in Poland." Journal of entrepreneurship, management and innovation 16, no. 1 (2020): 169-193.
[23]. Kiran, T., and A. Reddy. "Critical success factors of ERP implementation in SMEs." Journal of Project Management 4, no. 4 (2019): 267-280.
[24]. Kapoor, Kawaljeet Kaur, Kuttimani Tamilmani, Nripendra P. Rana, Pushp Patil, Yogesh K. Dwivedi, and Sridhar Nerur. "Advances in social media research: Past, present and future." Information Systems Frontiers 20, no. 3 (2018): 531-558.
[25]. Tadesse, Michael M., Hongfei Lin, Bo Xu, and Liang Yang. "Personality predictions based on user behavior on the facebook social media platform." IEEE Access 6 (2018): 61959-61969.
[26]. Bhaskar, Hari Lal. "Business process reengineering framework and methodology: a critical study." International Journal of Services and Operations Management 29, no. 4 (2018): 527-556.
[27]. Eshuis, Rik, and Pieter Van Gorp. "Synthesizing data-centric models from business process models." Computing 98, no. 4 (2016): 345-373.
[28]. AbdEllatif, Mahmoud, Marwa Salah Farhan, and Naglaa Saeed Shehata. "Overcoming business process reengineering obstacles using ontology-based knowledge map methodology." Future Computing and Informatics Journal 3, no. 1 (2018): 7-28.

[29]. Bhaskar, Hari Lal, and R. P. Singh. "Business process reengineering: a recent review." Bhaskar, HL, and Singh, RP (2014). Business process reengineering: a recent review. Global Journal of Business Management 8, no. 2 (2014): 24-51.

[30]. Dinçkan, Meltem. "Business Process Reengineering via Simulation Technique and a Case Study." PhD diss., Marmara Universitesi (Turkey), 2013.

[31]. Li, Hui, Qi Chen, Zhaoman Zhong, Rongrong Gong, and Guokai Han. "E-word of mouth sentiment analysis for user behavior studies." Information Processing & Management 59, no. 1 (2021): 102784.

[32]. https://en.wikipedia.org/wiki/Python_(programming_language).

[33]. Mansournia, Mohammad Ali, Angelika Geroldinger, Sander Greenland, and Georg Heinze. "Separation in logistic regression: causes, consequences, and control." American journal of epidemiology 187, no. 4 (2018): 864-870.

[34]. Hengl, Tomislav, Madlene Nussbaum, Marvin N. Wright, Gerard BM Heuvelink, and Benedikt Gräler. "Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables." PeerJ 6 (2018): e5518.

[35]. Kumar, Sachin, Aditya Sharma, B. Kartheek Reddy, Shreyas Sachan, Vaibhav Jain, and Jagvinder Singh. "An intelligent model based on integrated inverse document frequency and multinomial Naive Bayes for current affairs news categorisation." International Journal of System Assurance Engineering and Management (2021): 1-15.

[36]. Anagaw, Amare, and Yang-Lang Chang. "A new complement naïve Bayesian approach for biomedical data classification." Journal of Ambient Intelligence and Humanized Computing 10, no. 10 (2019): 3889-3897.

[37]. Chaudhary, Kiran, Mansaf Alam, Mabrook S. Al-Rakhami, and Abdu Gumaei. "Machine learning-based mathematical modelling for prediction of social media consumer behavior using big data analytics." Journal of Big Data 8, no. 1 (2021): 1-20.