Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 3s (2025)

Nonlinear PESTLE Framework-Based Sentiment Analysis Using Machine Learning for Root Cause Identification

Sudarshan S. Sonawane
Department of Computer Engineering,
R. C. Patel Institute of Technology, Shirpur, India
sudars2000@gmail.com

Article History:
***Received**: 23-09-2024*
***Revised**: 02-11-2024*
***Accepted**: 19-11-2024*

Abstract:
Sentiment analysis has gained significance in understanding the intricate dynamics of public sentiment in contemporary social media and news communication. This research proposes a nonlinear sentiment analysis model that integrates the PESTLE framework with machine learning techniques to identify and quantify the influence of external factors—political, economic, social, technological, legal, and environmental—on sentiment patterns. The model employs a Support Vector Machine (SVM) classifier with nonlinear kernel transformations to capture complex relationships between sentiment triggers and influencing factors. Experimental results demonstrate the efficacy of the nonlinear approach in correlating sentiment classifications with their intrinsic influences, offering a more robust understanding of root causes. Future work will explore advanced nonlinear classifiers and additional external factors for enhanced real-world applicability.

Keywords: PESTLE, SVM classifier, Machine Learning System, Real-Time Environment.

1 Introduction

Globally many of the events, actions or causes are either the outcome of a sentiment, and in vicious circle, the events or actions lead to certain sentiments. Advent developments in the predictive models are enabling the entities to focus on the dynamics of sentiments, understand how the sentiments are emerging or influencing the happenings around.

Today, with the rapid penetration of social media solutions, scope for farther reach of one’s expression is phenomenal. Unlike yesteryears, wherein an individual has limited medium to voice their opinions, today there are seamless set of social mediums wherein the individuals can express their viewpoints, at almost no cost and it can reach to millions of people, based on what level of content-sharing networks it penetrates in. Once a viewpoint is expressed and it reaches certain section of viral groups, there is no stoppage or control mechanism as to how it can influence the viewer’s mindset [1].

Influencing the mindsets, views, and decisions of people thru public mediums is decades old practice. Right from small group gatherings to contemporary social media platforms, across the medium the practice of influencing the sentiments or sentiments leading to action is prevalent. For instance, any mishap or a crime taking place certainly could trigger news and sentiment. However, what nurtures that sentiment possibly could be the technology element. If there is repeated sharing of such information from one group to the other, it leads to some prejudice formation on such occurrences [2].

In last decade of time, there are many novel studies about the sentiment analysis topic. Right from the medium impact to the topics that influence the sentiments, and mechanisms to analyze the sentiments, there are numerous solutions that were discussed in the process [3]. However, the rapid change in the quantum of data available, assertive changes to the consumer trends, and other such factors are creating space for paradigm shift in how the sentiment analysis is perceived [4].

Majority of the studies that is available in the literature refers to the outline on subject of sentiment analysis, or the techniques and tools for carrying out the facets of sentiment analysis. There is need for more intrinsic understanding of the issue, as to exploring the cause-and-effect analysis. In this manuscript, the attempt is to develop an internal mechanism that can be a plug-in element to understand the cause-and-effect element in the Sentiment Analysis spectrum.

1.1 Gap Analysis

Sentiment analysis is seen as a source of information, wherein some underlying insights could be interpreted for better decision making. For instance, analysts using the social media communication exchange data for analysis, is about exploring the disparity or neutrality analysis. This is quite prevalent in the case of brand or product related perception, more aggressive analysis on understanding how the sentiments are prevailing pertaining to electoral process, which parties has positive outlook or negative outlook etc. Similarly, any small or big global or regional events too trigger the sentiments of people [5].

With social media being an accessible medium of expression, in general it has become a potential source of data capture which leads to more intrinsic analysis. Majority of the techniques deployed in the sentiment analysis relies on the assessment of combination of good words to event words or occurrence and try assimilating the sentiment disparity [3] [6]. However, one critical dimension that is amiss in the current analysis models are about understanding the difference between the actual event trigger and the factors influencing it.

In an illustrative case scenario, for instance if there is gradual rise in the crude oil prices, thus impacting the fuel prices at national level, certainly this occurrence could be triggering some negative sentiment. But the media consistently playing this news or more coverage to such a news, combined with economic instability of an individual, leads to more negativity about the incident. In simple terms, if the actual cause leads to x quantum of negativity, the other encompassing factors leads the negativity to X^4 levels [7]. Hence, in addition to understanding where and how the issues are prevailing on the sentiment conditions, there is need for exploring the proportions that impact the overall process outcome. Thus, in this manuscript, a solution is discussed wherein a popular management theory of PESTLE analysis is proposed for Sentiment Analysis modelling.

1.2 Objectives

To identify the root problem for a specific trigger and identifying the correlation element that profoundly influences the sentiments.

Create a systematic modeling that can support in effective analysis of the sentiments, coordinated models that can assist in understanding the factors influencing the sentiment more.

1.3 Organizing of the Manuscript

In the section-2 of this manuscript, summarized understanding of the related work is presented. Section-3 reflects more on the actual model discussed, the materials and methods, process flow and algorithm approach for the model. Section-4 provides insights into the experimental analysis conducted for the proposed solution. Section-5 offers conclusion on the model, and the interpretation from the experimental outcome.

2 Related Work

Across the industrial and academic research studies, one of the significant topics explored in the data science domain is the scope of sentiment analysis. Since last decade, there are distinct set of solutions explored in the domain. There are versatile range of sentiment analysis models imperative from the literature. Some of the key aspects which are seen as significant in the process of developing a sentiment analysis model is discussed in this section.

Strings, Words and Semantics are standing as the basic elements, around which any comprehensive set of sentiment analysis models are evolving. Indexing, screening the posts for positive words, adjectives, and certain set of words integral to referring the emotions, moods, mindset, or perception of the users are seen as the critical functionality integral to classifying the emotions. Across the spectrum of research papers reviewed, there are patterns and combination of words, phrases or emotions are used as the preliminary structure for models [6-11].

Followed by, the other integral factor of performing the sentiment analysis models are about sources and stream of data being adapted by the analysis engines. For instance, in the case of exit poll kind of analysis, the models might rely on sample data from the post-event aspects, but in the case of an live event, there is phenomenal importance of streaming data, data sources and the pace in which the analysis can be executed by the process [12-14].

The other key classification evident in the system is about how the machine learning kind of intelligence systems are being adapted in the model. The comprehensive review literature studies indicate the potential ways in which the genetic algorithms, and distinct set of classifiers can be potential use for managing the process of classifying or testing the sentiments across the topics. While some models were created agnostic to any specific topic or activity, few of the sentiment analysis models are attributed to specific kind of social media tools. The complexities in the sentiment analysis models are envisaged in the case of unstructured data and comprehensive data, unlike the scale-metric data conditions [15-19].

The studies have focused on intrinsic issues like the quality of index words or screening process, scope of classifiers and time, performance quality in capturing the right kind of sentiments. But there are possibly very few studies that targeted the underlying factors impacting the sentiment analysis, and the gap is wide for future study explorations.

3 PESTLE Framework based Sentiment Analysis Patten

3.1 Rationale for the Model

In the general management theory, the external factors related analysis for a product is carried out based different theoretic models, and one such popular framework for analysis is the PESTLE analysis model.

In the PESTLE framework, the fundamental element is analyzed using different aspects like the political, economic, social, technological, legal, or environmental. A simpler understanding of the subject is how any event or occurrence or news has the influence of the mentioned elements. By adapting such detailed framework analysis, it helps in understanding where and how the influence is, and the measures essential for addressing such elements to boost or neutralize the facets. Applying the same to the context of sentiment analysis modeling and spectrum, though the root-cause issue for a sentiment arising is one element, the actual boosting of such sentiment could be a triggered by other elements. Identifying specific elements or the weightage of influence by each of the elements towards a specific problem can lead to more intrinsic analysis, and better decision-making process. Aiming at the scope for such analysis, it is presumed that the model proposed in this manuscript can be a potential solution [20].

3.2 PESTLE Framework

An organisation, business, product, as well as service can have its outcomes considerably impacted by external variables such as Political, Sociological, Economic, Technological, Legal, as well as Environmental elements that are studied inside a PESTLE analysis. The model's simplicity in applying to many circumstances is one of its main advantages, so this framework has been utilised for decades inside the method of making strategic decisions.

One of the underlying motivations for performing sentiment analysis was related to the decision-making process, as well as the PESTLE system, if applied towards the analysis itself, will aid in the identification of specific types or macro-environmental elements [21].

Investigate the primary environmental influences on an organisation (political, sociological, economic, technological, legal, as well as environmental). It may be applied in a variety of situations and help senior managers as well as human resources experts make strategic decisions. It can be used as an organisational tool to recognize the elements that cause development or decline in relation to problems, potential, as well as operational direction, as complement to the aspects of the company position [22] Image 1.

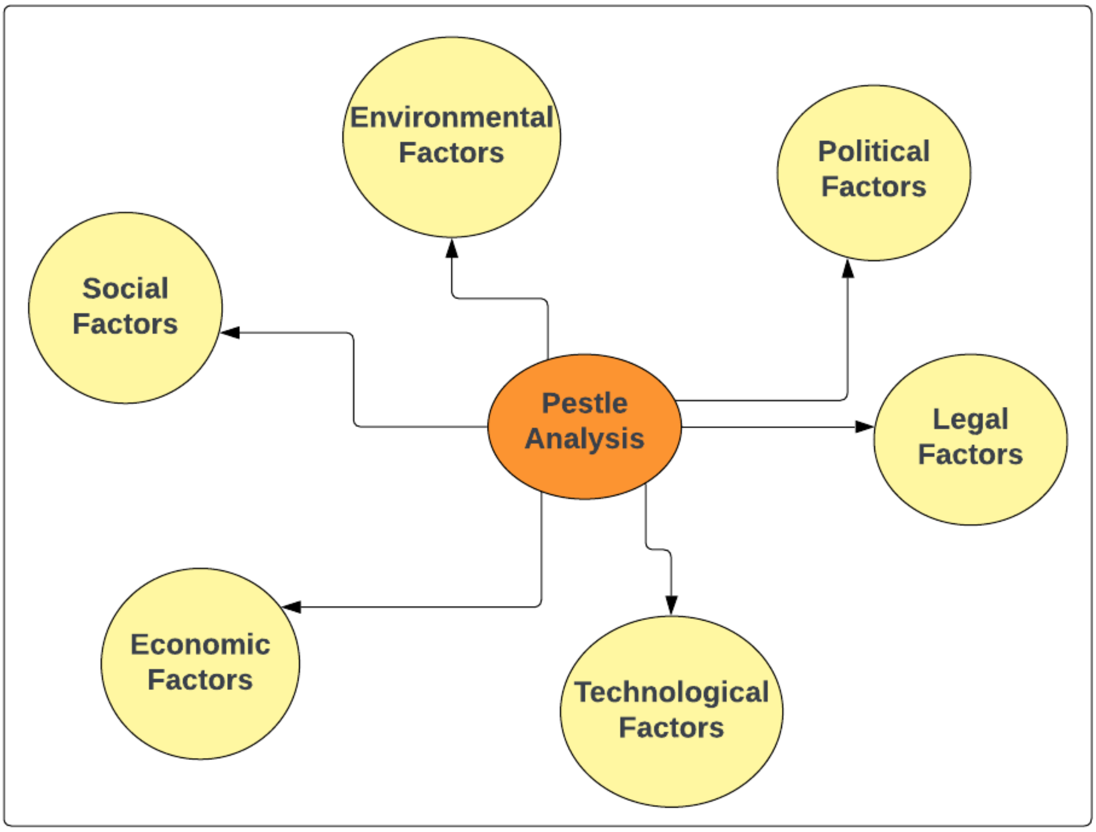

Figure 1: PESTLE Analysis

3.3 Model Narrative

Sentiment analysis using the PESTLE framework pattern is explained in this sub-section.

Firstly, for a news or a viewpoint or an occurrence, using any of the existing solutions, the positive or negative or neutral sentiment view related outcome is adjudged. Post the process of adjudgment, the application of PESTLE framework shall be initiated to understand the correlation of the chosen external factors influencing the current sentiment pattern.

In the application of the model, the process adapted is to have the weightage mechanism for each of the six factors (Political, Economic, Sociological, Technological, Legal and Environmental) and use them for detailed analysis. The preform followed for the weightage process is mentioned in the table below Table 1.

Table 1: The weightage process

External Factor	**Classification Criteria Description**
Political	The tweets or views or posts garnered should have words or strings related to politics.

	The tweets or views or posts garnered can be from political affiliated sources like politicians, parliamentary committees, or governance related parties. For each of such view or post, one point is added to the rank volume defined as RV. In the given period, total no. of RVs for a factor is seen as important metric.
Economical	The tweets or views or posts garnered should have words or strings related to economics. The tweets or views or posts garnered can be from affiliated sources like economists, finance professionals, banking professionals, parliamentary committees, or governance related parties. For each of such view or post attributing to economic dimension of words or posts, one point is added to the rank volume defined as RV. In the given period, total no. of RVs for a factor is seen as important.
Social	The tweets or views or posts garnered should have words or strings related to social aspects like culture, religion, society, socio-culture etc. The tweets or views or posts garnered can be from political affiliated sources like activists, NGOs, or welfare societies, charity people, philanthropies, public views. For each of such view or post, one point is added to the rank volume defined as RV. In the given period, total no. of RVs for a factor is seen as important.
Technology	The tweets or views or posts garnered should be viral in large scales and the volume of forwards or likes or re-tweets shall be the index criteria for adding the rank volume points. The tweets or views or posts garnered gathering more forwards or engagements or likes or dislikes etc. are the criteria for the rank volume addition. For each of such view or retweet or forwards or group messages, one point is added to the rank volume defined as RV. In the given period, total no. of RVs for a factor is seen as important.
Legal	The tweets or views or posts garnered should have words or strings related to Legal, para-legal terminologies. The tweets or views or posts garnered can be from affiliated sources like legal teams, law firms, advisory firms, market research firms, paralegal teams etc. For each of such view or post attributing to economic dimension of words or posts, one point is added to the rank volume defined as RV. In the given period, total no. of RVs for a factor is seen as important.
Environmental	The tweets or views or posts garnered should have words or strings related to environmental, climate, weather, earth, air, water etc. The tweets or views or posts garnered can be from affiliated sources like activists, NGOs, agro communities, farming groups, etc.

	For each of such view or post attributing to economic dimension of words or posts, one point is added to the rank volume defined as RV. In the given period, total no. of RVs for a factor is seen as important.

- **Indexing Pattern for Words or Strings**

The indexing for each of the external factor discussed shall be an addition of words felt appropriate to the context, and each of the word could be added to multiple external factors among the six chosen factors. In the case of one word being in multiple factor index, still the ranking volume points are added to the individual factor as independent [23].

For example, if there is a news on election, and view posts are having words like "poor policies", poor policies could be a string in economic, legal or political. In such cases, all the three factors will garner one respective point addition to the ranking volume.

- **Indexing the Post Sources**

Affiliated sources could be a manual addition, or a filtered source detection model adapted to add or revise to the existing source database in the system. Once the system captures such information, such sources are considered as authentic [19].

- **Ranking Volume Approach**

For each of the post as positive or negative or neutral, there is Y number of tweets that are used as the basis. And that Y is used as the total impact factor records (record is the cumulative number of posts or views or tweets related to specific topic). The SIX external factors from PESTLE are seen as independent factors influencing the "Y" which is a dependent factor outcome.

Using the statistical modelling solution, the correlation values are calculated between the dependent factor "Y" and the other "X" factors. Higher the correlation between the Y and X factors shall be interpreted as high impact factor or low impact or moderate impact factor.

By understanding the underlying influence factor, the decision teams can work towards making appropriate decision making, which can help in accomplishing the task more effectively.

Computation of X and Y force Range Score between 1-5 and Clustering

For both X and Y factors the range score is defined between 1-5, wherein the X_1 between certain volume can be ranked as 1, and bigger range can be ranked as 2 and there on.

{example: Ranking Volume of 100-1000 can be scored as 1, and Ranking volume of 10001-3000 can be ranked as 2, and Ranking Volume of 8000 or above can be ranked as 5}

For each of the topic or event or occurrence, there could be no. of groups or hashtags or mediums in which the topic is discussed, thus, each of them can be grouped into one cluster.

{illustrative scenario: on a topic called "$MeToo$", there could many hashtags, and groups where multiple people could be discussing. Thus, each of them could be a cluster defined as C, and hence, the matrix computation for the sentiment is based on $C_1 ... C_n$

Machine Learning System

The aforementioned process is executed using a supervised machine learning model wherein the whole process is attributed to a specific systematic screening process. The classifier used for the process is the SVM classifier wherein the hyperplane pattern is used for classifying the information or a post as positive or negative. Similarly, for the external factors analysis, the indexing holds critical value, and for each of the six factors, the index shall be developed, and it shall be updated on periodical basis with more appropriate words or modifying or deleting the existing words or strings [24-26].

3.4 Process Flow

The following process flow refers to the critical steps followed in the process of estimating the underlying impact factor in the system.

3.4.1 Process Flow Procedural Description

Step 1. Update the ML system for index directory of words for each of the six external forces $(PESTLE)$ seen as independent variable factors.

Step 2. Update the ML system for sources classification for index directory for each of the six external forces $(PESTLE)$ seen as independent variable factors.

Step 3. Choose the Post or the Topic for Analysis

Step 4. Using the pre-chosen Sentiment Analysis system, identify the post or event related sentiment as negative or positive or neutral

Step 5. The cumulative number of tweets, posts, views, or opinions garnered for the step-4 shall be used as the ranking volume (RV) for Y, thus it is denoted as $Y(RV)$

Step 6. Using the index directory, for each of the six external factors the $X(RV)$ values are identified.

Step 7. Based on the inputs used, the range are estimated, and accordingly the rank score for both X and Y factors are developed.

Step 8. Following the cluster model, the no. of mediums or groups or hashtags are chosen as independent cluster C and the process of Step-4 to step-7 is followed.

Correlation Analysis is conducted for the dependent and independent variables using regression analysis.

The top-2 high correlation values shall be seen as underlying trigger factors influencing the decision. The least-2 correlation values shall be presumed to be having negligible influence on the decision.

Step 9. Train the Machine Learning Classifier using the above pattern, with trained dataset for supervised learning process

Step 10. Decision making teams can work their strategic decisions based on the inputs garnered from the system.

3.5 Algorithm Flow

Begin

{

Let I_w be the index dictionary of words for each of the external force classification

Let I_s be the index dictionary of affiliate sources for each of the external force classification

Let "XYZ"be the systematic approach followed for sentiment value assessment

Analysis of Trigger Factor

{

Computation of Y {sentiment result} Ranking volume

$$Y = \sum(Posts + Tweets + Opinions + etc)$$

Computation of X forces Ranking Volume

For each of the X force,

$$X_1 = \sum(Matching\,words + Matching\,sources)...\{X_1\ can\ be\ denoted\ for\ Political\}$$

$$X_2 = \sum(Matching\,words + Matching\,sources)...\{X_2\ can\ be\ denoted\ for\ Economical\}$$

$$X_3 = \sum(Matching\,words + Matching\,sources)...\{X_3\ can\ be\ denoted\ for\ Social\}$$

$$X_4 = \sum(Matching\,words + Matching\,sources)...\{X_1\ can\ be\ denoted\ for\ \mathrm{Technical}\}$$

$$X_5 = \sum(Matching\,words + Matching\,sources)...\{X_1\ can\ be\ denoted\ for\ \mathrm{Legal}\}$$

$$X_6 = \sum(Matching\,words + Matching\,sources)...\{X_1\ can\ be\ denoted\ for\ Environmental\}$$

Final Ranking Volume Matrix shall be in the form of Matrix below Table 2

Table 2: Ranking Volume Matrix

Cluster (C)	C1	C2	Cn
Y {Sentiment View}	Y(RV)	Y(RV)	Y(RV)
X1	X1(RV)	X1(RV)	X1(RV)
X2	X2(RV)	X2(RV)	X2(RV)
X3	X3(RV)	X3(RV)	X3(RV)
X4	X4(RV)	X4(RV)	X4(RV)
X5	X5(RV)	X5(RV)	X5(RV)

X6	X6(RV)	X6(RV)	X6(RV)

Range Score for the above matrix Table 3

Table 3: Range Score for the above matrix

Cluster (C)	C1	C2	Cn
Y {Sentiment View}	(1-5) range score	(1-5) range score	(1-5) range score
X1	(1-5) range score	(1-5) range score	(1-5) range score
X2	(1-5) range score	(1-5) range score	(1-5) range score
X3	(1-5) range score	(1-5) range score	(1-5) range score
X4	(1-5) range score	(1-5) range score	(1-5) range score
X5	(1-5) range score	(1-5) range score	(1-5) range score
X6	(1-5) range score	(1-5) range score	(1-5) range score

Analysis of Regression and Correlation

We all are aware that the interaction between two variables is modelled using linear regression. So, the following is a straightforward regression analysis equation:

$$Y = a + b\left(X\right)$$

Where,

$$Y \;=\; Dependent\;\; variable$$

$$X \;=\; Independent\;\; variable$$

$$a = \frac{\left[\left(\Sigma(y)\Sigma(x2)\right) - \left(\Sigma(x)\Sigma(xy)\right)\right]}{\left[n\Sigma(x2) - \Sigma(x)2\right]}$$

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 3s (2025)

$$b = \frac{\left[\left(n\Sigma(xy)\right) - \left(\Sigma(x)\Sigma(y)\right)\right]}{\left[n\Sigma(x2) - \Sigma(x)2\right]}$$

Convergence Coefficient

The equation for such line of linear regression is as follows:

$$Y = b_0 + b_1(X)$$

b_0 is a constant in this instance, while b_1 is indeed the regression coefficient.

The following is a formula for such regression coefficient.

$$b1 = \frac{\sum(yi - y)(xi - x)}{\sum(xi - x)^2}$$

x_i as well as y_i, which represent the observed data sets, x as well as y, which represents this mean value of each variable, respectively,

We are aware that there regression analysis equations as well as 2 regression coefficients.
The formula for said regression coefficient between y as well as x is:

$$byx = \left(\frac{\sigma y}{\sigma x}\right) r$$

A formula of x on y regression coefficient is:

$$bxy = \left(\frac{\sigma x}{\sigma y}\right) r$$

Where,

σx Denotes deviation of x

σy Denotes deviation of y

The following is a listing of a few characteristics of such a regression coefficient.

- This regression coefficient was represented by the number b. byx can be used to indicate this regression coefficient from y on x. bxy can be used to indicate the regression coefficient between x 3 and y. Each of the two regression coefficients will be less than 1 if one is bigger than 1.

- • They are not unaffected by the shift in scale. If x and y were multiplied either by constant, then regression coefficient would change for both of them.
- Its arithmetic mean between both regression coefficients exceeds or if it is equal to a correlation coefficient.
- The correlation coefficient is the geometric mean of the two regression coefficients.

byx also was positive if bxy was positive, as well as vice versa.

Visit us here at BYJU'S - Its Learning App to learn much about correlation as well as regression formulae, the distinction between the two, as well as examples.

- This box of regression formulae is pasted from the public domain and needs to be customized for this manuscript.

Interpretation Validations

- The top-2 high correlations from six of the X factors to Y, shall be seen as high influence for the current sentiment
- The least-2 correlations from six of the X factors to Y, shall be seen as negligible influence for the current sentiment
- The middle-2 range correlations from six of the X factors to Y, shall be seen as moderate influence for the current sentiment

4 Experimental Study

The data analysis model discussed in the earlier section was trained into the SVM classifier, and the model is tested based on 350 set of data records worked manually for sentimental data classification and used the pattern for the analysis. The sentimental data and its internal analysis process is assessed, and the structured data alongside the interpretation is used for the analysis purpose.

To detail the iteration on how each of those 350 is assessed for the influence category, one such dataset related to political topic is compiled and depicted below, where in it reflects on the systematic approach of classification followed. In one iteration, a global view on the # Metoo topic is used for analysis. Using the 12 different groups for analysis, the patterns were modelled Table 4.

Table 4: Using the 12 different groups for analysis

Y (Sentiment)	**X1 (Environment)**	**X2 (Economic)**	**X3 (Environment**	**X4 (Technological)**	**X5 (Legal)**	**X6 (Social)**
4	5	3	5	5	2	3
3	3	2	4	1	5	1
1	5	3	2	4	5	4

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 3s (2025)

3	3	1	1	3	5	3
3	3	4	4	5	3	2
2	4	3	3	3	2	1
3	4	3	1	5	3	5
5	4	1	3	1	4	5
4	5	5	2	5	3	4
5	5	4	4	1	4	1
2	1	2	3	5	3	2
5	4	2	4	1	3	4

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.793769284							
R Square	0.630069676							
Adjusted R Square	0.186153288							
Standard Error	0.977575233							
Observations	12							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	6	8.1384	1.3564	1.41934313	0.358903886			
Residual	5	4.778267	0.955653					

Total	11	12.91667						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	1.776645461	2.326899	0.763525	0.479607402	-4.204837701	7.758129	-4.20484	7.758129
X1	0.018808707	0.258123	0.072867	0.944736934	-0.644718002	0.682335	-0.64472	0.682335
X2	-0.841451253	0.32541	-2.58582	0.04908459	-1.677943206	-0.00496	-1.67794	-0.00496
X3	0.236851908	0.519264	0.45613	0.66742458	-1.097959924	1.571664	-1.09796	1.571664
X4	0.171073417	0.263827	0.648431	0.545304562	-0.507115214	0.849262	-0.50712	0.849262
X5	0.376850925	0.247789	1.520854	0.188777063	-0.260111018	1.013813	-0.26011	1.013813
X6	0.551326701	0.311563	1.769552	0.137026654	-0.249571129	1.352225	-0.24957	1.352225

As per the interpretation validations, the Social and Legal factors have more direct underlying influence, and the technological aspects stand moderate, while economic factors stand the lower levels of influence for the sentiments envisaged for the process.

- Note : the model do not show any disparity or judgment to whether the sentiment is right or wrong or the degree of its correctness. The model shall attempt to identifying the possible underlying factors that is influencing the resulting sentiment.

Machine Learning Classification Training Outcome.

Total of 350 decision records on the sentiment views and its relative inherent influence factors analyzed data were used for training the machine learning models. Based on the data collated, the SVM classifier-based system is trained for the study. The table 5 represents the quantum of data used for training and the quantum of data used for testing the model.

Table 5: the quantum of data used for training and the quantum of data used for testing the model

Total Dataset	350
Training Dataset	210
Testing Dataset	140

Based on the interpretation of the results generated, the following figurative representation along with the data labels is depicted, indicating the basic performance elements of the proposed system Figure 2.

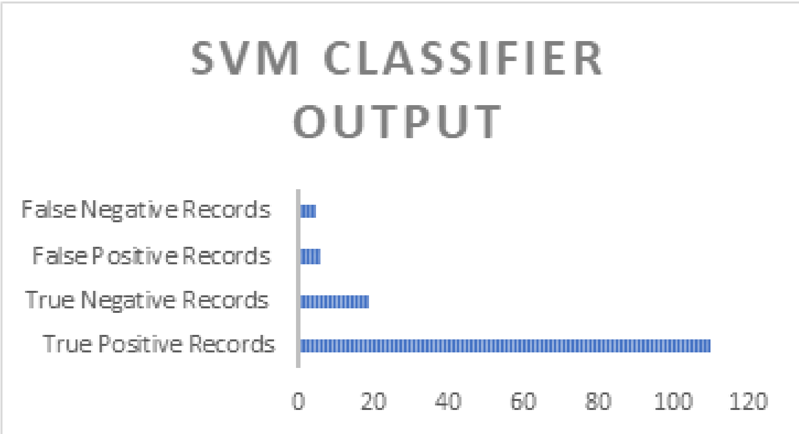

Figure 2: the data labels is depicted, indicating the basic performance elements of the proposed system.

The success of applying the machine learning models can be envisioned based on the performance metrics like the accuracy, sensitivity of the system and other such metrics that can signify the efficacy of the system. The outcome of true and false records generated by the SVM classifier trained model is computed for the metrics. The performance of the classifier based on the computation is depicted in the graphical representation below figure 3.

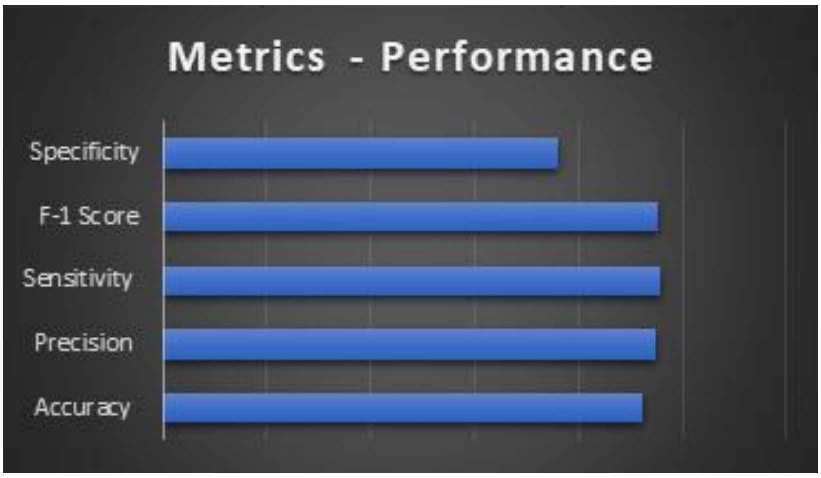

Figure 3: The performance of the classifier based on the computation

If the system can be trained in more effective ways, and more classification of interpretation is carried out in the future studies, any discrepancy in terms of poor performance of the classifier or the model pertaining to any index related issues could be observed and fine-tuned. In general, based on the fundamental pattern analysis, it is evident that the system is resourceful and is working effectively.

5 Conclusion

Sentiment Analysis as a systematic approach is gaining importance, and globally there are many research studies undertaken in fine-tuning the process of sentiment analysis. Multiple dimensions are explored in the case of assessing the sentiments related to distinct topics, events, news or brands etc. In this manuscript, a novel attempt is made in terms of creating a system that can be resourceful in understanding the trigger factors that could be instrumental for the actual sentiment observed in the process. As a preliminary factor, using the management model of PESTLE analysis the attempt is made for classifying the inherent influencing factors. The model is trained over the SVM classifier, and the outcome has generated potential outcome in the system. While the model seems to be significant in the performance (based on the result interpretations), for the future studies, there are multiple dimensions in which the system could be explored. Firstly, application of different set of external or internal influence factors as inherent influencers, and the other way by exploring the use of genetic algorithm approaches to improve the reinforced learning practices. If the system can be implemented for real-time environment, there could be more potential insights emerging in the process.

References

[1] Fang, Xing, and Justin Zhan. "Sentiment analysis using product review data." Journal of Big Data 2.1 (2015): 1-14.
[2] Karim, Mirsa, and Smija Das. "Sentiment analysis on textual reviews." IOP Conference Series: Materials Science and Engineering. Vol. 396. No. 1. IOP Publishing, 2018.
[3] Drus, Zulfadzli, and Haliyana Khalid. "Sentiment analysis in social media and its application: Systematic literature review." Procedia Computer Science 161 (2019): 707-714.
[4] Chandra, Rohitash, and Aswin Krishna. "COVID-19 sentiment analysis via deep learning during the rise of novel cases." Plos one 16.8 (2021): e0255615.
[5] Iglesias, Carlos A., and Antonio Moreno. "Sentiment analysis for social media." Applied Sciences 9.23 (2019): 5037.
[6] Keith Norambuena, Brian, Exequiel Fuentes Lettura, and Claudio Meneses Villegas. "Sentiment analysis and opinion mining applied to scientific paper reviews." Intelligent data analysis 23.1 (2019): 191-214.
[7] Rani, Sujata, and Parteek Kumar. "Aspect-based Sentiment Analysis using Dependency Parsing." Transactions on Asian and Low-Resource Language Information Processing 21.3 (2021): 1-19.
[8] Ligthart, Alexander, Cagatay Catal, and Bedir Tekinerdogan. "Systematic reviews in sentiment analysis: a tertiary study." Artificial Intelligence Review 54.7 (2021): 4997-5053.
[9] https://www.kdnuggets.com/2020/06/5-essential-papers-sentiment-analysis.html
[10] Zhou, Jin, and Jun-min Ye. "Sentiment analysis in education research: a review of journal publications." Interactive learning environments (2020): 1-13.
[11] Puschmann, Cornelius, and Alison Powell. "Turning words into consumer preferences: How sentiment analysis is framed in research and the news media." Social Media+ Society 4.3 (2018): 2056305118797724.

[12] Mäntylä, Mika V., Daniel Graziotin, and Miikka Kuutila. "The evolution of sentiment analysis—A review of research topics, venues, and top cited papers." Computer Science Review 27 (2018): 16-32.
[13] Alsaeedi, Abdullah, and Mohammad Zubair Khan. "A study on sentiment analysis techniques of Twitter data." International Journal of Advanced Computer Science and Applications 10.2 (2019): 361-374.
[14] Puteh, Nurnasran. "Sentiment Analysis with Deep Learning: A Bibliometric Review." Turkish Journal of Computer and Mathematics Education (TURCOMAT) 12.3 (2021): 1509-1519.
[15] Swati, Ubale, Chilekar Pranali, and Sonkamble Pragati. "Sentiment analysis of news articles using machine learning approach." International journal of advances in electronics and computer science 2.4 (2015): 114-116.
[16] Shi, Yong, et al. "Survey on classic and latest textual sentiment analysis articles and techniques." International Journal of Information Technology & Decision Making 18.04 (2019): 1243-1287.
[17] Parthasarathy, G., and D. C. Tomar. "Sentiment analyzer: Analysis of journal citations from citation databases." 2014 5th international conference-confluence the next generation information technology summit (confluence). IEEE, 2014, pp. 923-928.
[18] Anvar Shathik, J., and K. Krishna Prasad. "A Literature Review on Application of Sentiment Analysis Using Machine Learning Techniques." International Journal of Applied Engineering and Management Letters (IJAEML), ISSN: 2581-7000.
[19] Ashir, Abubakar M. "A Generalized Method for Sentiment Analysis across Different Sources." Applied Computational Intelligence and Soft Computing 2021 (2021).
[20] Buye, Ronald. "Critical examination of the PESTLE Analysis Model." Action Research for Development: Research Gate (2021).
[21] Sammut-Bonnici, Tanya, and David Galea. "PEST analysis." (2014).
[22] Matović, Ivana Marinović. "PESTEL Analysis of External Environment as a Success Factor of Startup Business." ConScienS Conference Proceedings. Scientia Moralitas Research Institute, 2020, pp. 96-102.
[23] Zunic, Anastazia, Padraig Corcoran, and Irena Spasic. "Sentiment analysis in health and well-being: systematic review." JMIR medical informatics 8.1 (2020): e16023.
[24] Liu, Bing, and Lei Zhang. "A survey of opinion mining and sentiment analysis." Mining text data. Springer, Boston, MA, 2012. 415-463.
[25] Cao, Jianfang, et al. "Improved support vector machine classification algorithm based on adaptive feature weight updating in the Hadoop cluster environment." PloS one 14.4 (2019): e0215136.
[26] Cervantes, Jair, et al. "A comprehensive survey on support vector machine classification: Applications, challenges and trends." Neurocomputing 408 (2020): 189-215.