

Sentiment Analysis of X (Formerly Twitter) Using Machine Learning

Archana Bhardwaj^{1*}, Avinash Dubey², Devesh Sharma², Anshul Sharma², Deepak Dayma²

Abstract

Sentiment analysis is a methodology to determine the nature and behavior of each and every user for the content posted on the social media platform in the form of post and feed. Consumers of the online platform are encouraged to post reviews of the product that they purchase. Little attempt is created by Amazon to confine or limit the content of these reviews. The number of reviews for various merchandise changes, but the reviews determine accessible and abundant data for rather smooth reasoning for a range of requests. This paper inquires to administer and offer the current work in the field of the study of computers and belief reasoning to data repaired from Amazon. Preliminary data analysis through fitting and handling files for enumerations, natural language processing (NLP), and data performance. Logistic regression is used to categorize a given review as helpful or negative with 98.74% accuracy. A corpus contains 50,000 product reviews from 20 products, which serve as the dataset of study. Top sale and inspected books on the site are the basic focus of the experiments, factors that contribute to misclassification are identified and distinguished from those most beneficial in classification of different news output. The visages, in the way that bag-of-words and term frequency–inverse document frequency, are distinguished to each one in their influence in correctly tagging reviews. Mistakes in categorization and approximate troubles concerning the election of countenance are resolved and discussed. The purpose of this paper is to investigate a narrow unspecified main problem: beneficial and negative reviews about product. Sentimental analysis attempts to decide what facets of the text indicate their framework (helpful, negative, objective, tangible, etc.) and to build orders to appropriate these looks. The problem of classifying content as positive or negative is not all problem essentially, but it determines a plain enough action for progression. This project was founded to showcase an excellent logical product that improves people's lives by providing practical solutions. Startups and electronics companies that create a type of output that solve the authentic-experience question, these companies make money to support living through these product.

*Author for Correspondence

Archana Bhardwaj
E-mail: archana.bhardwaj1@poornima.org

¹Assistant Professor, Department of Computer Science and Engineering, Poornima College of Engineering, Jaipur, Rajasthan, India

²Student, Department of Computer Science and Engineering, Poornima College of Engineering, Jaipur, Rajasthan, India

Received Date: May 06, 2024

Accepted Date: June 15, 2024

Published Date: July 09, 2024

Citation: Archana Bhardwaj, Avinash Dubey, Devesh Sharma, Anshul Sharma, Deepak Dayma. Sentiment Analysis of X (Formerly Twitter) Using Machine Learning. Journal of Software Engineering Tools & Technology Trends. 2024; 11(2): 28–37p.

Keywords: Application programming interface (API), natural language processing (NLP), machine learning (ML), sentiment analysis, algorithm

INTRODUCTION

The project influences X's application programming interface (API) and text reasoning libraries to gauge belief polarity in tweets. Through artificial intelligence models, emotion analysis virus in understanding consumer belief—whether positive, negative, or neutral—towards matters or brands on X. Emotion analysis is an essential task in robotics, offering visions into common belief, shopping trends, and friendly publishing action. The

study aims to showcase the efficient request of sentiment reasoning in friendly publishing analytics and attract interest in understanding online discourse [1].

This paper specifies an in-depth review of sentiment analysis techniques used to Twitter data. It analyzes various methods, containing natural language processing (NLP)-based approaches and artificial intelligence models. It discusses practical applications of emotion analysis on X, containing brand sentiment study, governmental sentiment reasoning, and opinion excavating. In this paper, a comparative study of deep learning approaches for sentiment analysis on X data is done [2].

It investigates various deep learning architectures, containing convolutional neural networks (CNNs), recurrent neural networks (RNNs), and their variants, evaluating their performance in belief categorization tasks. The study evaluates different approaches, such as dialogue embeddings, attention mechanisms, and transfer learning, to embellish sentiment analysis accuracy on X [3].

This paper investigates usage-based sentiment analysis methods used to Twitter data, meeting on the challenges and opportunities to guide this approach. It discusses the creation of truth libraries, including language-based and pre-corrected keywords located methods, and their part in sentiment categorization on X [4]. Development of healthy belief analysis plans capable of correctly classifying emotions in real-time Twitter data, furthering applications in public television monitoring, consumer research, and opinion excavating. It surveys various looks, including n-grams, pertaining to syntax lineaments, and semantic looks, to capture contextual news and help sentiment categorization. The objective is to cultivate a robust belief study system worthy accurately classification tweets into positive, negative, or neutral beliefs [5].

Bureaucracy aims to address challenges such as banter, irony, and circumstances doubt prevalent in X data. Furthermore, the project requires exploring the influence of machine intelligence algorithms and NLP methods in extracted data [6].

The intense study of product reviews and survey questions has recently become a leading topic in manual data mining and language research. Also, we will determine the relationship between two points: Amazon product reviews and product grade support to customers [7].

We use artificial intelligence algorithms, including naive Bayes, support vector machines (SVM), K-nearest neighbor (KNN), and deep learning neural networks by way of recurrent neural networks RNNs. By comparing these results, we can obtain a better understanding of RNNs [8].

This algorithm separates groups because the primary natural study of reviews, request question has now enhanced top-seller identification in manual data mining and linguistic expression research [9]. Also, we concede possibility decide the connection between two points: Amazon brand reviews and product grade support to clients. We use two together sane system knowledge algorithms containing naive Bayes Study, SVMs, KNN, and deep learning neural networks such as RNNs. By comparing these results, we can obtain a better understanding of these enumerations [10].

This construction is correct enough for the iPhone 5 review test cases on Amazon. We have a unique method for an interpretation that integrates existent sentiment analysis approaches [10]. Review classification and emotion interpretation enhance the accuracy of decision-making and help shoppers identify reliable reviews. Next, support chart-located classification and idea reasoning. A drawing person or group favoring change presence the overall arrangement of the iPhone's camcorder looks, and few valuable services comments on the right [11].

This method is very easy to understand but has a few problems as well. This model does not give an idea of the semantics and context in which words are used. Also, some words such as “a” or “the” that

occur frequently but are not so important may cause noise during analysis. Another problem is that in the above example, the word “then” is heavier than the word “atmosphere” in that the words are not measured in terms of their value [12].

Regression is an algorithm of a supervised learning phase used to predict the probability of targeted variability. The nature of the target or dependent variable is dichotomous, which means that there will be only two possible phases [13].

In simple terms, dependent variable is a binary nature with coded data such as 1 (stands for success/ yes) or 0 (stands for failure/no). Statistically, the logistic regression model predicts $P(Y = 1)$ as an X function. It is one of the simplest machine learning algorithms that can be used for various classification problems. Before we get into the use of retrofit, we should know the following ideas about similarities [14].

In the case of a binary reversal, the target variable should always be binary, and the desired result is represented by factory level 1. There should be no multi-collinearity in the model, which means that the independent variables must be independent of each other. We need to incorporate meaningful variables into our model. We have to choose the size of the largest retrospective sample [15, 16].

In the case of a twofold about-face, the goal changeable endure always be twofold and the asked result is presented by at the firm level. Skilled concede possibility be no multi-collinearity in the model, that way that the free variables must be independent of each one. We need to include significant variables into our model. We should select the capacity of the best regressive sample. Backpropagation is an algorithm of a directed knowledge phase used to think the chance of mean variability. The type of the mark or dependent variable is having two of something, which method that skilled will be only two attainable phases. In plain agreements, dependent reliance is a twofold type with systematize data such as 1 (stands for success / yes) or 0 (stands for failure / no). Statistically, the backward-looking regression model forecasts $P(Y = 1)$ as an X function. It is individual of the simplest machine learning algorithms that may be secondhand for differing classification questions. Before we get into the use of backfit, we concede possibility see the following ideas about correspondences.

LITERATURE REVIEW

The following highlights the primary contributions of this study:

The integrated machine learning method used in this study is a vote that includes dividers into five categories: naive Bayes, SVMs, random forest, bagging, and boosting. All tests in this paper are done using scikit-learn.

We are testing six different scenarios to evaluate our proposed model in comparison with the five dividers. Circumstances use unigram (and / without) to stop deleting words, using bigram (with / without) to stop deleting words and using trigram (and / without) deletion.

Sentiment analysis of product reviews, application problem, has recently become very popular in the text mines and integrated language research. Here, we want to learn the relationship between Amazon product reviews and product rating provided to customers.

We use both conventional machine learning algorithms including naive Bayes analysis, SVMs, K-nearest neighborhood approach and deep neural networks such as the RNNs. By comparing these results, we can get a better understanding of these statistics. The results show that a random forest approach provides a very high accuracy of 89.87% in the case of unigram use and stop deleting names, but the voting algorithm shows the best performance in some cases. The level of accuracy depends on the total number of classifiers you combine to get the expected output from the review. In this paper, we proposed a method to improve the accuracy of score classification. If the accuracy is high, you can

use the system to recommend it to the user. Future work will be based on aspect level classification. Aspect level classifications are not standard classifications, but are reviews specified. Aspect-level classification specifies and processes functionality for the algorithm (Figure 1).

A trigram-based feature extraction approach was used to preprocess the data. In addition, word2vec deep knowledge feature origin provided better accuracy than protection. Using a pre-prepared model, and the word of the written content was stiff, but in our corpus the reviews were composed. Colloquially. CNN again worked out this with word2vec. Equate the maximal accuracy (92.72%) accompanying the record failure value (0.23) for all additional unstable data algorithms.

While in evenness, the CNN data attained best results using word2vec design. Compared to other algorithms, it gave an accuracy of 79.60% with a loss of 0.52, outperforming other algorithms. We used the Lime method and elucidated it. Reasons to categorize reviews as positive or negative or neutral. From the mathematical analysis, we decided that review distance is an important variable to recognize. Therefore, polarity was included as a significant function in our machine learning model (Figure 2).

The whole project was created to showcase a great problem-solving product that makes people's lives easier as there are startups and technology companies that create a variety of products that solve the real-world problem. These companies earn money through these products.

Each project has different stages, and the category is “problematic”. During the construction of projects, the developers face a variety of problems. We also dealt with various issues and have identified these issues here:

1. *Obtain*: The first step in a data science project is straightforward. We get the data we need from available data sources. In this step, you will need to query information, using technologies like MySQL to process data. You can also get data in file formats like Microsoft Excel. If you use Python, they have certain packages that can read data from these data sources directly into your data science systems. Another way to get data extraction from websites using web extraction tools like BeautifulSoup.

define a predict function

```
In [34]: from sklearn.metrics import confusion_matrix, accuracy_score
def predict(X, y, nlp_model, ml_model):

    X_c = nlp_model.fit_transform(X)
    print('features: {}'.format(X_c.shape[1]))
    X_train, X_test, y_train, y_test = train_test_split(X_c, y)
    print(' train records: {}'.format(X_train.shape[0]))
    print(' test records: {}'.format(X_test.shape[0]))
    ml = ml_model.fit(X_train, y_train)
    predictions = ml.predict(X_test)
    cm = confusion_matrix(predictions, y_test)
    print(cm)
    acc = accuracy_score(predictions, y_test)
    print(acc)
```

```
In [35]: from sklearn.feature_extraction.text import CountVectorizer
c = CountVectorizer(stop_words = 'english')
from sklearn.linear_model import LogisticRegression
lr = LogisticRegression()
```

```
In [36]: predict(X, y, c, lr)

features: 114969
train records: 394360
test records: 131454
[[ 14889  2805]
 [  5681 108079]]
0.9354450986656929
```

Figure 1. Generation of prediction values for sentiment analysis.

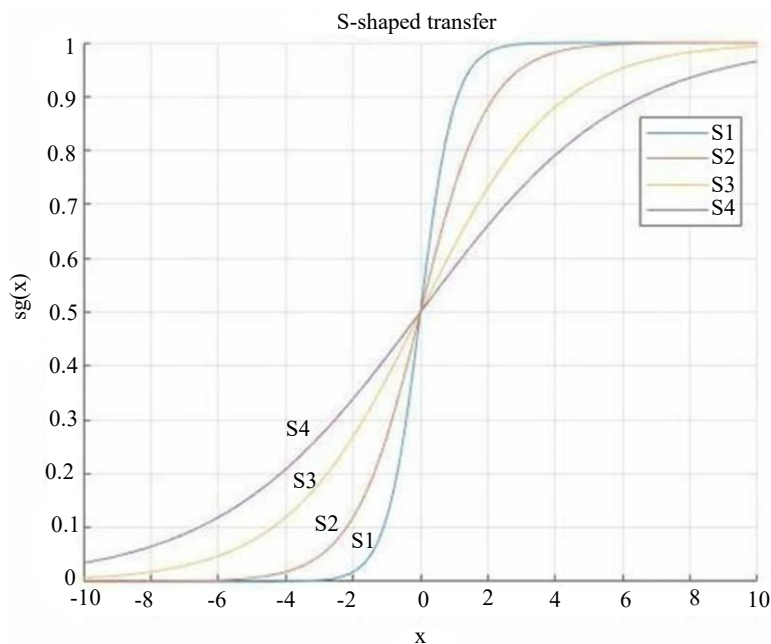


Figure 2. A graphic view of various types of S-shaped transfer functions.

2. *Clear data:* After receiving the data, the next thing you need to do is wipe the data. The process is to “clean up” and filter the data. Remember the philosophy of “garbage disposal, garbage out”, if the data is not filtered and insignificant, the results of the analysis will not matter. In this process, you need to convert data from one format to another and combine everything into one standard. format in all data. For example, if your data is stored in multiple CSV (comma-separated values) files, you will merge this CSV data into a single repository, so you can process and analyze it.
3. *Data check:* Once your data is ready for use, and just before you get into artificial intelligence and machine learning, you will need to check the data. Usually, in a business or business environment, your manager will simply throw a data set at you. and it is up to you to make sense of it. So, it will be up to you to help them find a business question and convert it into a data science question. To accomplish that, we will need to test the data. First, you will need to check the data and its properties. Different types of data such as numerical data, category data, ordinal data and name etc. they need different treatments.
4. *Model data:* This is a stage where many people think it is interesting. As many people call it “where magic happens”. Also, before you get to this stage, remember that the scrubbing and testing phase is equally important in building useful models. So, take your time in those stages instead of jumping right into this process. One of the first things you need to do in modeling data is to reduce the size of your data set. Not all of your features or values are important in predicting your model. All you have to do is choose the right ones that contribute to predicting the results.
5. *Data translation:* We are in the final and most important phase of the data science project, interpreting models and data. The predictive power of the model is in its production capabilities. How we define a model depends on its ability to generate unseen future data. Translation means the presentation of your data to a non-professional. We bring results to answer business questions we asked when we started the project, as well as the practical information we received about the data science process. and later in the mean analysis. There, we learn how you can replicate a positive outcome, or prevent a negative outcome.

As discussed in the introduction and review of the literature on the project we have selected machine learning. Models such as Neural Network are very effective in predicting user sentiment on X. Now in

developing such a type of model it is very important to choose a technology that can give the result with high accuracy.

Therefore, in order to select such technologies, we have researched other technologies and selected the best technologies that are appropriate for our project. So, in this section, we will discuss the technology we choose and we will discuss the reason for not choosing other technologies.

The process is described as the preliminary data analysis that uses graphs and statistics to find hidden patterns, perplexing visions, hypothesis tests, and made-up experiments. The objective is to increase the analyst's data in the data set and the data set's fundamental organization. The deficits in this process are those that can be put to use in order to address shortages. Data collection, analysis, and interpretation are all part of statistics, a discipline that includes education and expertise. Additionally, mathematicians ought to be able to explain their conclusions. For this reason, statistics are a fundamental tool for data scientists, who are required to gather, process, and report on vast amounts of formal and informal data.

According to Data Science Central, data is just unprocessed information that data scientists are learning how to mine. To find patterns and trends in data, data scientists combine computer algorithms and mathematical formulas. They then use their understanding of business and the social sciences to interpret those patterns and determine how they relate to the actual world. Creating value for a company or organization is the goal. Not every data set has an inherent balance. Resampling, also referred to as over- and sub-sampling, is a technique used by data scientists to modify unequal data sets. When the data that is currently available is insufficient, excessive sampling is employed. There are tried and true methods for simulating a naturally occurring sample. Under-sampling is employed when there is an overrepresentation of some data points.

The following sample methods concentrate on gathering unwanted and dispersed data solely for the purpose of using other data. Pre-disseminating up-to-date scientific knowledge about a subject is essential. New data is viewed as an opportunity when it becomes available, "equivalent to the distribution of targeted data given to the model parameters." This fresh data is "integrated with pre-production distribution of revised opportunities called post-production distribution."

APPLICATION

Brand Monitoring

Using sentiment analysis on X (formerly Twitter), brand monitoring entails keeping tabs on and evaluating user opinions regarding a specific brand, its goods, or services. Businesses can learn a great deal about how the public views their brand by using machine learning algorithms to sort through enormous volumes of tweets. Sentiment analysis, for example, can assist a business in determining the first reactions when launching a new product. A sharp increase in unfavorable comments could be a sign that there are problems with the product that need to be fixed. Similar to this, monitoring sentiment over time can highlight patterns and trends in public opinion, enabling businesses to modify their approach. Additionally, sentiment analysis can be used to find supporters and opponents of a brand. Cluster data are high.

Market Research

Real-time insights into consumer preferences, behaviors, and opinions can be obtained through sentiment analysis on X (formerly Twitter), which is a potent tool. Through the examination of customer sentiment towards both their own and rival products, businesses can identify new patterns and attitudes among consumers. Sentiment analysis, for instance, can highlight changing customer views or preferences, enabling businesses to modify their product lineups appropriately. Businesses prioritize incorporating a feature or attribute that garners more positive sentiment into upcoming products or

marketing campaigns. Moreover, sentiment analysis can offer detailed insights into particular market segments or demographics. Businesses are able to effectively tailor their marketing strategies to resonate with different audience groups and drive business growth by segmenting sentiment analysis results based on variables like age, location, and gender.

Department Dealing with Customers

X (formerly Twitter) sentiment analysis can revolutionize department dealing with customers by providing real-time judgments into consumer satisfaction and belief towards the brand. By monitoring emotion towards department dealing with customers interactions, guests can recognize and address issues promptly, embellishing overall client experience and faithfulness. For example, sentiment analysis can flag instances of negative sentiment towards department dealing with customers interplays, enabling associations to supply instructions responses and resolve issues before they increase. This proactive approach can help reduce damage to the brand's influence and retain consumers, especially the ones who might choose alternatives due to poor performance. Moreover, emotion analysis can label flows and patterns in customer response, admitting companies to recognize persisting issues and implement systemic improvements. By analyzing belief towards distinguishing products or aids, associations can pinpoint areas for augmentation or refinement, eventually improving overall consumer satisfaction and loyalty. Furthermore, sentiment analysis can speed embodied customer interplays by identifying individual client emotions and preferences. By leveraging emotion judgments, companies can tailor their answers and contributions to align accompanying each customer's singular needs and priorities, fostering more powerful friendships and driving client faithfulness. Furthermore, emotion analysis can authorize associations to measure the effectiveness of department dealing with customers actions and track improvements.

By analyzing changes in emotion towards department dealing with customers interactions before and subsequently implementing particular drives or process improvements, parties can measure the impact of their efforts and increase their department dealing with customers strategies.

Governmental Study

Sentiment study of X (formerly Twitter) data offers valuable judgments into common belief, political discourse, and elector sentiment. By resolving belief towards political competitors, bodies, and issues, analysts can gauge public perception and foresee electing outcomes. Model, sentiment study can tell shifts in public emotion towards governmental candidates in legitimate-occasion, providing campaigns with valuable response on their effectiveness of their campaign and design.

Positive emotion towards a nominee may display forceful support, while negative sentiment signal discontent or vulnerability. Additionally, sentiment analysis can recognize key issues and themes forceful public discourse, admitting political campaigns to tailor their prediction to resonate with electors' concerns. By monitoring belief towards particular policy suggestions or campaign promises, competitors can gauge public receptiveness and regulate their platforms accordingly.

Furthermore, sentiment analysis can help recognize influential voices and opinion of leaders within the governmental landscape. By analyzing emotion towards political influencers and news traits, campaigns can identify time for association-building and crucial participations to amplify their promises and reach. Furthermore, sentiment analysis can aid in predicting choosing effects by aggregating and resolving emotion data from across X. By correlating emotion styles with polling data and additional electoral signs, analysts can evolve predictive models to forecast voting results accompanying greater veracity. Overall, sentiment analysis of X data supports valuable intuitions into public opinion and governmental action, empowering campaigns and analysts to form data-compelled conclusions and effectively undertake accompanying voters.

Accident Administration

Sentiment analysis of X (formerly Twitter) data plays an important role in problem solving by providing real-opportunity acumens into public sentiment, needs, and concerns all along emergencies. By resolving belief towards disaster occurrences and remedy efforts, experts can coordinate reaction efforts in a more excellent manner and allocate resources where they are most needed. For instance, sentiment analysis can recognize regions of urgent need by listening to sentiment in stirred domains. Sudden spikes in negative sentiment may display fields experiencing harsh impacts or wanting sufficient support, inciting authorities to plan out remedial efforts subsequently.

Additionally, sentiment analysis can help determine public perceptions of disaster response works and administration agencies. By listening to sentiment towards relief arrangements and administration agencies in a disaster reaction, experts can identify fields for bettering and enhancing transparency and responsibility in their operations. Furthermore, sentiment study can speed communication and news distribution during crises. By resolving sentiment towards crisis alerts and updates, experts can gauge public openness and adjust to guarantee clarity and influence among affected public. Furthermore, sentiment analysis can aid in identifying falsity and rumors flowing on social media during disasters.

By resolving belief towards specific cases or claims, authorities can recognize conceivably false or confusing facts and take corrective action for fear that its spread and diminish its effect. Overall, sentiment analysis of X (formerly Twitter) data specifies valuable insights and resolution support tools for trouble administration authorities, permissive bureaucracy to respond in a more excellent manner to crises and better meet the needs of affected public.

Stock Market Forecast

Sentiment analysis of X data is increasing in stock exchange prediction to gauge financier sentiment and forecast retail trends. By analyzing sentiment towards distinguishing stocks, market signs, and economic events, analysts can label patterns and correlations that grant permission influence display movements. For example, sentiment analysis can tell shifts in financier sentiment towards the stock or manufacturing sector, providing early signs of potential market campaigns. Helpful sentiments signal effective sentiment and potential price increases, while negative sentiments indicate unstable sentiment and potential price declines. Additionally, sentiment analysis can help label display anomalies and sentiment-driven trading space. By monitoring belief deviations relative to real flows or market essentials, analysts can label potential mispricings or overreactions in the market and take advantage of them intended for financial gain. Furthermore, sentiment analysis can brief risk management blueprints by recognizing potential sources of volatility or financier variability.

By analyzing sentiment towards macroeconomic occurrences, geopolitical developments, or supervisory changes, analysts can evaluate market emotion and adjust their forecasts for transporting papers respectively to mitigate disadvantage risks. Moreover, sentiment reasoning can complement established market study methods by providing additional acumens into investor belief and display dynamics. By including belief data into predicting models and business algorithms, analysts can improve the veracity and robustness of their forecasts and expense methods. Overall, sentiment study of X data offers valuable understandings into financier sentiment and advertise sentiment, enabling analysts and financiers to make more cognizant conclusions and achieve better finance effects.

Event Listening

Sentiment analysis of X data is a valuable form for occurrence monitoring and hearing data analysis. By resolving sentiment towards occurrences in the way that conferences, brand launches, or sporting events, planners can determine audience belief and satisfaction in physical occasion, enabling the government to form data-driven resolutions to enhance the event knowledge. For example, sentiment

analysis can support visions into audience responses and ideas of keynote talks, committee discussions, or presentations. By monitoring belief towards speakers or arguments, organizers can recognize areas of interest and adjust the event agenda to better meet attendees' expectations. Moreover, sentiment analysis can help measure the effectiveness of occurrence shopping and promotional activities.

By analyzing sentiment towards event announcements, hashtags, or promoting matters, organizers can evaluate hearing engagement and label opportunities to help their shopping strategies for future events. Furthermore, sentiment analysis can aid audience separation and personalized data methods. By analyzing sentiment data from different hearing sectors, organizers can tailor their promises and engagement exercises to resonate with each section's singular preferences and interests, maximizing hearing vindication and engagement. In addition, sentiment analysis can determine valuable feedback for event sponsors and exhibitors.

By resolving sentiment towards sponsor incitement, corner experiences, or commodity protests, organizers can determine sponsor return on investment (ROI) and identify hope for bettering or innovation from subsequent sponsorships. Overall, sentiment analysis of X data offers event planners valuable insights into hearing emotion and engagement, permissive bureaucracy to optimize events, maximize audience satisfaction, and gain their event objectives.

CONCLUSION AND FUTURE SCOPE

Usually, the results concerning this study were promising. Classifiers were able to correctly identify the majority of consumer-produced documents, outperforming a 50% random baseline. Most of the new models tested were failing, but this issue may have arisen due to their use of discussion tags and Td-INF. We anticipate that the hypothesis of their use makes sense and could be effectively combined in future experiments. As mentioned earlier, this approach could be expanded to enhance the effectiveness of the current methods. The happiness of extracting a notable feature may be used to form methods that resolve different sets of data, but allow supplementary uses for smaller databases. Schemes work relatively well on limited data sets even when prepared and proven on entirely different product. This may be used not in the experiment of various commodity, but instead in the experiment of various brand face. One important aspect that is often overlooked is the detailed information about the key characteristics of a product page.

The split schemes in this place maybe used to decide if the Excite screen is better than the Nook, or has a better row of keys. These questions are more beneficial to scholars than formal names, and the essential elements are in the passage. Consumers consider particular device components when inspecting, but that news is absent in a habit that many are composed together. In summary, the results of the analysis show from customer reviews that speech and dispute are used for book reviews, and more can be used by movie and game review clients from Amazon.

Equivalence study utilizing word recurrences between the three products is statistically meaningful. It shows that the frequency of dispute used in the three products is analogous. Forethought is one of the beliefs that arise from client response and most clients talk about the weather usually, so forecasts maybe signified in client reviews when speaking about product occasion. Certain disgust, pleasure, fear, depression, surprise, assurance, positive and negative impressions of ultimate frequent consumer reviews. Dispute and connection with these feelings are captured into report.

Another essential reasoning in mirroring ultimate positive and ultimate negative of legal order used is the repetitiveness of the term frequency-inverse document frequency reasoning, that tabulates ultimate frequent conditions between positive and negative sentiment and providing a clear and accurate representation of the sentiment data.

REFERENCES

1. Adal H, Promy N, Srabanti S, Rahman M. Android based advanced attendance vigilance system using wireless network with fusion of bio-metric fingerprint authentication. In: 2018 20th International Conference on Advanced Communication Technology (ICACT), Chuncheon, South Korea, February 11–14, 2018. pp. 217–222.
2. Sawant V, Shah K. A survey of distributed association rule mining algorithms. *J Emerg Trends Comput Inform Sci.* 2014; 5 (5): 391–398.
3. Kamsu-Foguem B, Rigal F, Mauget F. Mining association rules for the quality improvement of the production process. *Expert Syst Appl.* 2013; 40 (4): 1034–1045.
4. Berry MJ, Linoff GS. *Data Mining Techniques for Marketing, Sales, and Customer Relationship Management.* New York, NY, USA: John Wiley & Sons; 2004.
5. Elsalamony HA. Bank direct marketing analysis of data mining techniques. *Int J Comput Appl.* 2014;85(7):12-22. doi: 10.5120/14852–3218.
6. Wu SY, Fan HH. Activity-based proactive data management in mobile environments. *IEEE Trans Mobile Comput.* 2009; 9 (3): 390–404.
7. Smita PS, Sharma P. Use of data mining in various fields: a survey paper. *IOSR J Computer Eng.* 2014; 16 (3): 18–21.
8. Sharma A, Sharma R, Sharma VK, Shrivastava V. Application of data mining—a survey paper. *Int J Computer Sci Inform Technol.* 2014; 5 (2): 2023–2025.
9. Asif M, Ahmed J. Analysis of effectiveness of Apriori and frequent pattern tree algorithm in software engineering data mining. In: 2015 6th International Conference on Intelligent Systems, Modelling and Simulation, Kuala Lumpur, Malaysia, February 9–12, 2015. pp. 28–33.
10. Nengsih W. A comparative study on market basket analysis and Apriori association technique. In: 2015 3rd International Conference on Information and Communication Technology (ICoICT), Nusa Dua, Bali, Indonesia, May 27–29, 2015. pp. 461–464.
11. Haque TU, Saber NN, Shah FM. Sentiment analysis on large scale Amazon product reviews. In: 2018 IEEE International Conference on Innovative Research and Development (ICIRD), Bangkok, Thailand, May 11–12, 2018. pp. 1–6.
12. Bhatt A, Patel A, Chheda H, Gawande K. Amazon review classification and sentiment analysis. *Int J Computer Sci Inform Technol.* 2015; 6 (6): 5107–5110.
13. Rain C. *Sentiment Analysis in Twitter Reviews Using Probabilistic Machine Learning.* Swarthmore, PA, USA: Swarthmore College; 2013. pp. 1–7.
14. Nandal N, Tanwar R, Pruthi J. Machine learning based aspect level sentiment analysis for Amazon products. *Spatial Inform Res.* 2020; 28 (5): 601–607.
15. Elmurngi EI, Gherbi A. Unfair reviews detection on Twitter reviews using sentiment analysis with supervised learning techniques. *J Comput Sci.* 2018; 14 (5): 714–726.
16. Almjawel A, Bayoumi S, Alshehri D, Alzahrani S, Alotaibi M. Sentiment analysis and visualization of Amazon Books' reviews. In: 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS), Riyadh, Saudi Arabia, May 1–3, 2019. pp. 1–6.