WILEY | Hindawi

Research Article

Ranking Analysis for Online Customer Reviews of Products Using Opinion Mining with Clustering

S. K. Lakshmanaprabu [iD],[1] **K. Shankar** [iD],[2] **Deepak Gupta**,[3] **Ashish Khanna**,[3]
Joel J. P. C. Rodrigues [iD],[4,5,6] **Plácido R. Pinheiro** [iD],[7] **and Victor Hugo C. de Albuquerque** [iD][7]

[1]Department of Electronics and Instrumentation Engineering, B. S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, India
[2]School of Computing, Kalasalingam Academy of Research and Education, Krishnankoil, India
[3]Maharaja Agrasen Institute of Technology, GGSIP University, Delhi, India
[4]National Institute of Telecommunications (Inatel), Santa Rita do Sapucaí, MG, Brazil
[5]Instituto de Telecomunicações, Lisbon, Portugal
[6]ITMO University, St. Petersburg, Russia
[7]Graduate Program in Applied Informatics, University of Fortaleza, Fortaleza, CE, Brazil

Correspondence should be addressed to Joel J. P. C. Rodrigues; joeljr@ieee.org

Sites for web-based shopping are winding up increasingly famous these days. Organizations are anxious to think about their client purchasing conduct to build their item deal. Internet shopping is a method for powerful exchange among cash and merchandise which is finished by end clients without investing a huge energy spam. The goal of this paper is to dissect the high-recommendation web-based business sites with the help of a collection strategy and a swarm-based improvement system. At first, the client surveys of the items from web-based business locales with a few features were gathered and, afterward, a fuzzy c-means (FCM) grouping strategy to group the features for a less demanding procedure was utilized. Also, the novelty of this work—the Dragonfly Algorithm (DA)—recognizes ideal features of the items in sites, and an advanced ideal feature-based positioning procedure will be directed to discover, at long last, which web-based business webpage is best and easy to understand. From the execution, the outcomes demonstrate the greatest exactness rate, that is, 94.56% compared with existing methods.

1. Introduction

The expression "web-based social networking" bunches up an extensive variety of online exercises, blogs, company exchange sheets, chats, service rating sites, microweb journals, and so forth [1]. In the present world, E-business destinations are the capital of the market. No one has the need to go outside the market due to the dependability of sites; these sites are more trustworthy to the point of being showcased [2]. Clients are absolutely looking toward E-commerce destinations while they sit tight for the items to go ahead. A large quantity of websites of settled and rumored organizations are propelling their items at this stage; keeping confidence in these sites is the extreme goal, and this is the issue [3]. Web shopping entries enable clients to purchase items by means of the web and furthermore get them convinced by the covered area for an ostensible charge, thus decreasing the time required to purchase an item [4]. These days, numerous clients like to experience these surveys to land at a sensible choice about the reasonableness of a specific item or administration according to their necessities [5]. Supposition mining takes in people's perspectives, tests, conduct, and sentiments toward individuals, people, issues, exercises, subjects, and

their highlights. The sentiment is extensive on the grounds that they are essential impacts of our practices [6]. The subject of motion pictures is of noteworthy enthusiasm among the long-range informal communication groups, perceived both by the immense number of people discussing motion pictures and additionally a critical distinction in their estimations [7]. Highlights-based conclusion investigation incorporates component extraction, feeling forecast, opining characterization, and discretionary outline modules [7].

Feature extraction recognizes those item perspectives which are being remarked by clients; an assumption forecast distinguishes the content containing supposition or assessment by choosing notion extremity as positive, negative, or unbiased [8]. The synopsis module feature is assessed by the clients expressly, utilizing accuracy as an assessment metric to approve the highlights extraction and investigation process [9] at a solitary snap to diagnose a huge number of reviews. Nowadays, soft computing methods are vigorously conveyed in E-commerce businesses as information warehousing, and "soft computing" is the core of information warehousing or of some other propelled innovations today [10]. Most of the information give basic reference to the moment when those other buyers bring about the site. On assumption, an examination and better-grained thought mining approach focuses on the subsequent highlights [10]. The subjective perspectives are an accumulation of feelings, audits, suggestions, remarks, evaluations, and individual experiences shared by various clients imparted through gatherings and informal organizations, alongside authentic information [11]. This freely open gathering of audits is a help for the reviewers as they get the opportunity to share and take in the distinctive parts of an item, or benefits like highlights, focal points, constraints, and providers [12]. The main contribution of the work to examine consumer satisfaction of items in E-commerce sites and improve the proposed demonstration separates the features from the item audits with the help of a DA advancement display. For grouping, the FCM features demonstrate considerably. Groups find the positioning ratings for a specific thing with subtle elements that rank last on the clients' opinions, and an ideal features-based positioning method considered. The structure of this manuscript is organized as the follows: Section 2 talks about the survey of the literature of opinion analysis. In Section 3, the examination of the motivation of this work is discussed. From that point, Section 4 talks about the proposed technique. Lastly, in Section 5, 6, and 7, simulation investigation results conclude our work with future extension.

2. Literature Review

Opinion mining alludes to the utilization of "natural language processing, computational linguistics, and text mining to distinguish" or group whether the motion picture is great or not based on message feeling. The support vector machine (SVM) by Basari et al. [13] is a set of regulated learning strategies that dissects information and perceives the examples that are utilized for grouping. This examination concerns with double orders which are characterized by two classes. Those classes are the positive and the negative. The positive class indicates great message opinion; and generally, the negative class demonstrates the unfortunate message supposition of specific films. This avocation depends on the precision level of SVM with the approval procedure utilizing a 10-fold cross-approval and perplexity grid. The hybrid particle swarm optimization (PSO) is utilized to enhance the decision of the best parameter with a specific end goal of taking care of the double optimization problem. The outcome demonstrates the change of the precision level from 71.87% to 77%.

In Singh and Dubey's study [14], opinions can be classified into constructive and aggressive, and their level can be measured based on the occasion (individuals, association, and social issues). Along these lines, it is fundamental that individuals assess the investigation of feelings and examinations toward any social issue, individuals, or substance. As of today, the greater part of investigation has been done on notion examination of items and administrations. In the investigation of occasions and issues, information is recovered from online networks like Twitter.

The proposed technique starts with content preprocessing of breaking surveys into words and evacuating stop words, trialed by content change for making watchwords and creating input vectors by Claypo and Jaiyen [15]. MRF including features selection is therefore widely adopted in choosing significant highlights from countless reviews. At that point, K-means is utilized for clustering into positive and negative reviews. From the test, MRF including determination can proficiently reduce the number of highlights in the informational collection, so computational time is totally reduced. Moreover, K-means can accomplish the best grouping execution with a contrasted and self-organizing map, fuzzy c-means, and hierarchical clustering. In this manner, the collaboration of K-means with MRF's highlights determination is a viable model for grouping Thai eatery surveys.

In Parashar and Gupta's study [16], the customer of an E-commerce site has no real way to evaluate the quality of an acknowledged item to peruse an enormous number of reviews. This exploration work centers on building up a basic leadership calculation which can assess the nature of an item by sorting past surveys on a size of numbers and showing it on an E-commerce site. Customers can make utilization of these positions furnished with items over any E-commerce site to settle on their own choices.

A novel swarm intelligence optimization strategy is proposed called the Dragonfly Algorithm (DA) by Mirjalili [17]. The proposed algorithm is benchmarked by few scientific test capacities and one genuine contextual investigation subjectively and quantitatively. The after effects of DA and BDA demonstrate that the proposed calculations can enhance the underlying irregular populace for a given issue, meet toward the worldwide ideal, and give extremely

focused outcomes contrasted with other surely understood calculations in the literature. The consequences of MODA additionally demonstrate that this algorithm tends to discover exceptionally precise approximations of the Pareto ideal arrangements with a high uniform dispersion for multi-objective problems.

Park and Kim [18] have proposed the application of the Jaccard distance score to web data mining. To address the exploration reason, information has been gathered from two distinct sources that mirror the perspectives of voyagers and co-op specialists. At that point, an arrangement of content information mining techniques was connected to distinguish the dialect contrasts among explorers and CVB sites, as indicated by the accompanying classes: shopping, feasting, nightlife/exercises, and attractions. Some conceivable methodological expansions that can enhance suggestion abilities and administrative ramifications of these discoveries are given.

Sentiment mining makes the procedure of determination and basic leadership simpler. In spite of the fact that few methods exist for the assessment of mining in view of basic leadership in this paper, Malika et al. [19] have proposed a novel approach notwithstanding the opinions produced from surveys gathered from E-commerce sites, illustrating the general notion for basic leadership. This has been fused as extra weights which can be entered by the client and balanced by the need. Motivation to do this is needed for a specific component of an item which may differ from individual to individual. Additionally, an official conclusion lies in the purchaser's part notwithstanding the feelings gathered and investigated from the reviews.

Xiao et al. [20] have recommended and proposed an estimation calculation, which coordinates identity characteristics with protection inclination power and afterward groups the clients according to identity attributes. Next, this paper accomplishes a cross-breed communitarian separating proposals by joining supposition examination with protection concern. Investigations demonstrate that this model can viably tackle the issue on any meager condition. All the more, critically, a blend of subjective protection concern and target suggestion innovation can diminish the impact of clients' security worries on their acknowledgment of the service.

3. Motivation for Study

(i) E-commerce has an enormous area to investigate research entries about everything like clicking of a client, routing of the client over the E-business site, and obtaining client comments.

(ii) The fundamental impediments are that we make clients rely on the coordination of words. So as an essential advance, we need to show to the clients how to utilize it.

(iii) The web-based life relates to the entire world and is one of the clarifications behind information overload

on the web. There are various structures in which a client delivers a product which is posted on the Internet.

(iv) The existing framework has given one component that anybody can give input about any item. The individual from testing E-shopping site can give counterfeit criticism to the first site [13, 21].

(v) Suppose the item have a great quality may get negative review due to the some cutomer erroneous entry. Hence the different clients avoid to purchase that item so, it become an ideal.

4. Methodology

Information mining is to remove data and learning which is not known by individuals and conceivably helpful from countless and unclear irregular information of useful application. In our paper, the new application situates model-proposed online items in various shopping sites in view of the audits of customers. At first, we created the extremity database of various items of distinctive sources from E-trade locales. Every item has its own highlights sets which are great markers in bunching the item audits based on a set of chosen features (attributes). All the user reviews refreshed amid the period are removed from the web, and a refreshed element-based synopsis is created. The whole client surveys which refreshed component-based outline are created for gathering and suing the FCM model. In the wake of collection examination, the DA optimization is considered to improve the features to discover ideal sites with items. Moreover, this investigation positioning procedure ranks the ideal highlights based on suggesting E-trade locales with RPN investigation. In the wake of handling client surveys from the above advances, an item is recommended in light of the general score.

4.1. Data Source. We prepared a crude dataset from online E-shopping sites (Amazon, Flipkart, Snapdeal, ShopClues and Paytm). Every one of these locales is extremely prominent in introducing the greater part from where the general population gets a kick-out of the chance to buy. These gathered surveys are identified via the administration of various products from those sites. Our investigation gathered an excess of customer reviews for various categories, and its subtle elements essentially depicted. In this module, the client selects an item from a given class. At first, we include portable items in the classification since cell phones are a standout among the most inspected and sold items on E-business destinations. The system appears in Figure 1.

From that, different item reviews having positive, negative, and some nonverbal importance remarks were checked for expelling any undesirable thing to play out any information mining usefulness. A large portion of the items includes things, and the greater portion of the words used to decide the extremity of these highlights
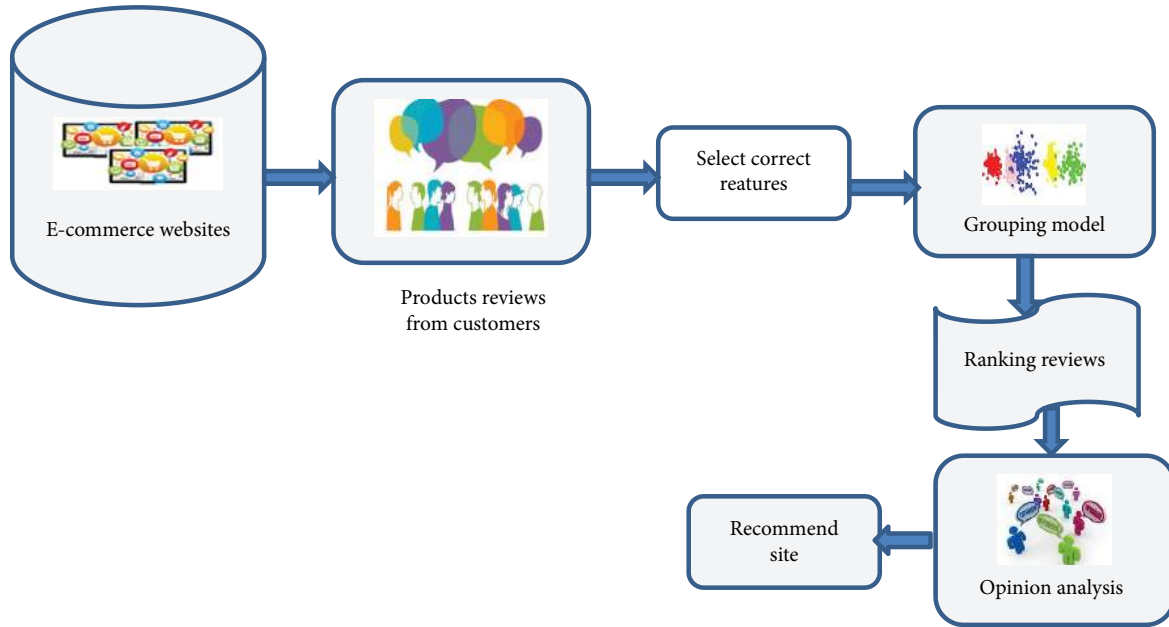
FIGURE 1: Diagram of the opinion mining process.

are descriptive words found in the region of the opinion feature.

4.2. Opinion Mining Analysis.
Opinions are a key to every single human action since they are key influencers of our practices. In reality, organizations and associations dependably need to discover client or general assessments about their items and administration. The machine learning procedure can enable us to know examples of the sentiments. From a business perspective, conclusion mining can enable the proprietors to comprehend their clients' needs and emotions through audits of items and administrations. Mostly this examination having three phases are feature extraction, grouping, and raking model. This feature is identified with a specific element taken into thought and constructed in light of those features. The whole execution abilities are tested and true, and furthermore, features proceed to decide a crisp or diverse arrangement of classes. The new of group of information in themselves and their evaluation is naturally or incrinsic.

4.3. Review Extraction.
The feature extraction and opinion identification is performed and features are separated from gathered item reviews from E commerce sites to enhance the quality of review for analysis. The connection among suppositions and items includes enhancement of the item audit rating. The example data attributes include cost, positive surveys, and quality, each included in every last audit demonstrating the client's articulations. The vast majority of trait data (features) are one of the qualities of items; the highlights from client surveys are utilized for scoring every item. Feelings on the web are broke down and looked at utilizing opinion eyewitness. This module is used to mine the opinion of customer reviews, summary of surveys, produce and store for additional step preparation.

4.4. Grouping Topology.
Since grouping of client opinions is helpful to different business perspectives, it has been a common strategy used to find many element articulations from content for a opinion mining application. Likeness measures utilized for bunching are typically in view of some type of distributional closeness. This paper proposes the collection of surveys from online destinations utilizing a fuzzy clustering method. Grouping is an unsupervised learning assignment, so no class esteems speaking to a previous blend of the information examples. It is generally utilized for taking care of grouping issues in different sorts of utilization. The reason for this grouping procedure is to recognize the clusters of information and allow an enrollment estimation of every datum example to each group [22].

4.4.1. Fuzzy Clustering.
The FCM is generally utilized for clustering, whereas the execution of the FCM depends on the determination of intital cluster head or cluster membership value to the features of reviews. It gives a strategy on how to assemble information that focuses on populating multidimensional space into a particular number of various clusters. The preferred fundamental standpoint of fuzzy c that implies grouping is that it permits continuous enrollments of information focused on clusters which are estimated as degrees. It figures out the group focus, utilizing Gaussian weights; utilizes expansive introductory models; and includes procedures for taking out bunching. The fundamental target of iterative bunching and fuzzy c-means calculation is to limit the weight inside grouping entirety of

squared blunder target capacity as shown in the following equation:

$$O_e = \sum_{i=1}^{d} \sum_{j=1}^{c} m_{ij}^{e} \left\| f_i - c_j \right\|^2, \tag{1}$$

where O_e is the objective function and fuzziness index and d, m, c is the membership of ith data to the jth cluster center, feature vector, and jth cluster center. The FCM enables each element vector to have a place with each bunch with a fluffy truth esteem (in the vicinity of 0 and 1), which is illustrated in (1). The calculation assigns an element vector to a bunch as indicated by the most extreme weight of the element vector over all the groups.

Fuzzy clustering by differentiating permits information that focuses on having a place within excess of one group. The subsequent segment along these lines is a fluffy parcel. Each clustering is related to an enrollment work that communicates how much direct individual information has a place in the group.

$$c_j = \left(\frac{\sum_{i=1}^{d} \left(m_{ij} \right)^{e} f_i}{\sum_{i=1}^{d} \left(m_{ij} \right)^{e}} \right). \tag{2}$$

The related cluster focuses on the structure of the information which is ideal as the calculation depends on the client to indicate the number of groups introduced in the arrangement of information to be grouped. At last, these means, with the exception of the underlying advances, are rehashed until the centroids never move again. From the technique of fuzzy clustering, the chosen data for groups are cost, quality, shipping charge, and a few other parameters. As per web-based social networking and client examination, the gathering of comparable stubborn individuals is called clustering [23]. It alludes to a technique by which the datasets of web clients are assembled into groups of little sets with comparative information.

4.5. Ranking Analysis. Features determination and clustering is based on ranking the item reviews enlivened, and one simple optimization is considered. The proposed framework is ranking items just chosen by the client, not all items. According to that, unique clusters locate the ranking rating for a particular item with data. At that point, the proposed framework look through the items with determination which is indicated by the client. This positioning in light of the base cost, maximum quality, optimal brand, and its target work is demonstrated as follows:

$$\text{Ranking} = \{\text{min\_cost, max\_quality, opt\_brand}\}. \tag{3}$$

To establish client trust, E-business people should set base criteria for the item; for example, if the reviews of an item are coming into this range, than they ought to be considered the further deal on their stage. The choice of ranking features model depends on the customer preference and updated the position of optimal features discussed in the following section.

4.6. Dragonfly Algorithm. Dragonflies are considered little predators that chase all other little bugs in nature. Fairy dragonflies likewise originate before other marine bugs and even little fishes. The fascinating certainty about dragonflies is their interesting and uncommon swarming conduct. These two swarming practices are fundamentally the same as the two primary periods of streamlining [24] utilizing metaheuristics: investigation and misuse. "New feature set based on objectives updating procedure having some default steps like separation, alignment, cohesion, attraction toward a food source; it's clearly discussed in the upcoming section."

4.6.1. Dragonfly Initialization. Initialize the population of dragonflies (here, the features from grouping) defined as D_i.

$$F_{e_i} = \left\{ F_{e_1}, F_{e_2}, F_{e_3}, \dots, F_{e_n} \right\}. \tag{4}$$

(1) Steps.

(i) Separation: the capacity of the separation process in DA is that it avoids the static smash of people from different people in the neighborhood. The separation S_e can be calculated by the following equation:

$$\text{Sep}_{ie} = \sum_{k=1}^{N} F_{e_i} - F_{e_k}, \tag{5}$$

where $Pe_i \rightarrow$ indicates the separation of the ith individual, F_e is the current position of the individual, F_{e_k} is the position of the kth individual, and N is the total number of adjoining individuals in the search space.

(ii) Alignment: similarly with that of dragonflies, this takes place based on the velocity matching of individuals to that of other individuals in the neighborhood. The alignment is calculated as

$$\text{Ali}_i = \frac{\sum_{k=1}^{N} E_k}{N}, \tag{6}$$

where $\text{Al}_i \rightarrow$ indicates the alignment of the ith neighboring individual, E_k is the velocity of the kth individual, and N is the total number of neighboring individuals in the search space.

(iii) Cohesion: cohesion means the tendency of individuals toward the center of the mass of the neighborhood. The cohesion Co_i (ith individual) can be calculated as

$$\text{Co}_i = \frac{\sum_{k=1}^{N} F_{e_k}}{N} - F_e. \tag{7}$$

(iv) Toward a food source: this is the attraction toward a food source among the dragonflies and the outward distraction of an enemy among the dragonflies:

$$\text{Food}_i = F_e^+ - F_e^-,$$
$$\text{Ene}_i = D^- + D, \tag{8}$$

where F_e^- indicates the position of the enemy and F_e^+ indicates the position of the food source.

(2) Updating Process.

$$\Delta \text{Feature\_set} = \left(w_s \text{Sep}_i + w_a \text{Ali}_i + w_c \text{Co}_i \right.$$
$$\left. + w_f \text{Food}_i + w_e \text{Ene}_i\right) \tag{9}$$
$$+ w\Delta \text{Feature}_t.$$

For updating the position of the artificial dragonflies in the scan space and for reenacting the developments of dragonflies, two vectors, specifically, the step ($\Delta Feature$) and position (*Feature*) vectors, are considered. Parameter documentation of that condition w_s, w_a, w_f, consists of weights of the compared practices, that is, detachment, arrangement, union, and factors. Subsequently, the enhanced positions are obtained from the yield of calculation and are additionally used to extricate the potential highlights by checking the score esteems with a limit.

4.7. Review Raking. After, we chose the optimal ranking opinions to assess the relative significance of each element as per related sentiment score and utilized that measure to rank highlights. Also, we trust that these commentators rating the sites have vital data that aids in distinguishing untruthful opinions. From that point, they recommended a perspective positioning calculation to rank the vital angles by thinking about both the viewpoint recurrence and the impact of opinions given to every perspective on their general sentiments.

5. Result Analysis

The proposed strategy was executed in a JAVA Platform Windows machine containing the following arrangement: Intel (R) Core i5 processor, 1.6 GHz, and 4 GB RAM. The example items with the site surveys are shown in Table 1.

Reviews with evaluations in this range represent almost 80% of aggregate surveys, mirroring for the most part a great involvement with the items on Amazon. The collaboration terms among goods and subject factors give us more data about how point factors influence accommodation votes. Table 1 demonstrates some of the most mainstream sites with real items; review scores are broke down. In database showcasing, through client database data, organizations can dissect the purchaser inclinations of clients and give clients distinguishing electronic lists to build the attraction list on customers.

Table 2 demonstrates the different product versus website scores esteem examination. In a large portion of the opinions, more users say, as regards some books, that substance of a book is unimportant, that the book appears to club everything together, and so forth. These sorts of irrelevance are arranged as an alternate component, and they are thought of negatively as an incentive to foresee the rank in more exact and precise measures. To examinations of the suggestion

TABLE 1: Numbers of review data.

Websites/ products	Electronic application	Mobile phone	Books	Music	Fashion
Amazon	155	229	134	114	67
Flipkart	135	134	67	105	119
Snapdeal	112	123	60	94	88
Paytm	28	45	21	48	30
ShopClues	33	25	69	25	78

TABLE 2: Score values of products with features in the proposed model.

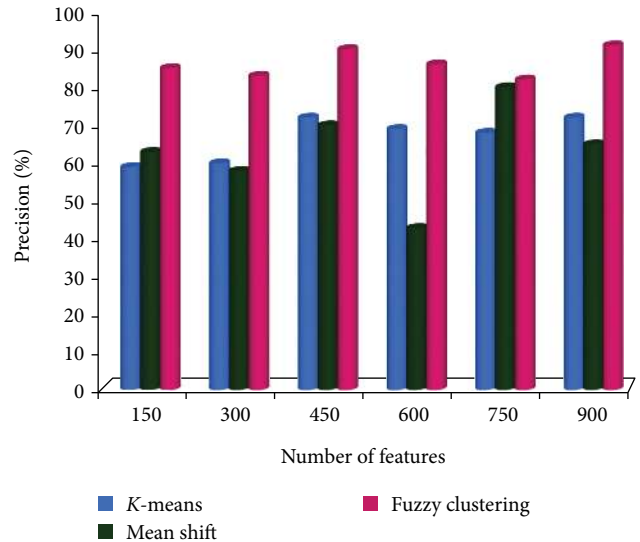| Products	Amazon	Flipkart	Snapdeal	Paytm	ShopClues
(P1) Electronic application	9.52	6.89	9.46	6.48	0.52
(P2) Mobile phone	10	8.55	9.2	7.45	0.19
(P3) Books	8.89	7.69	8.55	6.78	0.45
(P4) Music	8.45	7.46	6.78	4.55	0.11
(P5) Fashion	9.55	8.59	6.55	7.54	0.38

FIGURE 2: Result of feature grouping.

framework, different sorts of proposal calculations are utilized. A few items on web-based business were taken in light of these sentiments. At this point, normal on exactness and normal on review is computed for each sentiment.

Figure 2 and Table 3 demonstrate all the techniques bunching the reviews into different groups. Each cluster incorporates either positive or negative comments. The trial after effects of fuzzy grouping contrasted the mean shift and K-means clustering in view of the features. The proposed bunching model's precision is at 84.62% for the audit information from 450 features, while contrasting with that of

TABLE 3: Results of the grouping model.

| Websites | Precision | | | Recall | | |
	Fuzzy	K-means	Mean shift	Fuzzy	K-means	Mean shift
Amazon	85.24	59.56	75.22	92.22	79.66	82.2
Flipkart	82.14	55.89	65.22	85.55	77.47	79.45
Snapdeal	76.55	69.45	61.2	93.21	79.45	65.5
Paytm	82.5	79.5	45.56	94.5	69.55	71.82
ShopClues	92.58	68.5	82.22	88.5	75.45	69.2
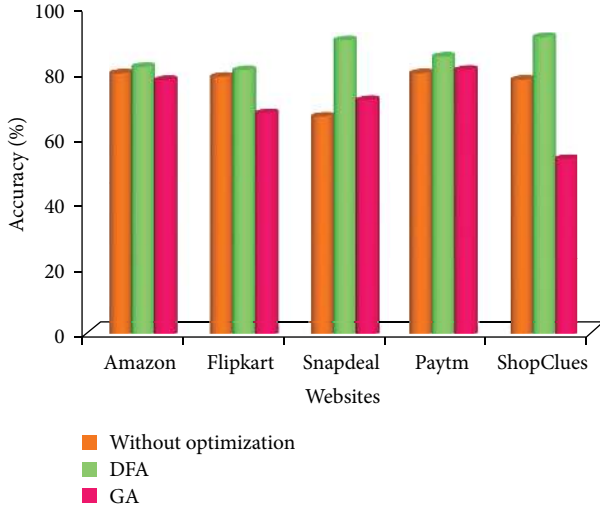
FIGURE 3: Optimization results.

other models. The graphical portrayal of a general examination of the execution measurements was obtained from the unthinkable esteems. The figure demonstrates the exactness and review of various sites in the wake of the grouping process—generally, an overall improvement of the measurements in our proposed model.

Figure 3 demonstrates the ranking of rating extremity of the audits, age of the review, and the helpfulness of item scores of the reviews. Our proposed optimization model can be compared to a model without optimization and genetic algorithm (GA). The general item illustration is done, and the best appraised item appears to the client with item investigation with dragonfly improvement, and those without enhancement are analyzed; among these examinations, the most extreme score rating goes to the Amazon site.

6. Performance and Ranking Analysis

To analyze the proposal framework, different kinds of suggestion algorithms are utilized. A few items on the Internet business were taken in view of assumptions. From that, normal on exactness and normal on review are illustrated for each opinion. Results illustrate (Figure 4) that the proposed structure gave the better outcome in the examination of the system. Sensitivity, specificity, precision, and F measure are low in the FCM examination when contrasted with the FCM with DA methods. FCM with a singular streamlining system gives better outcomes contrasted with FCM;
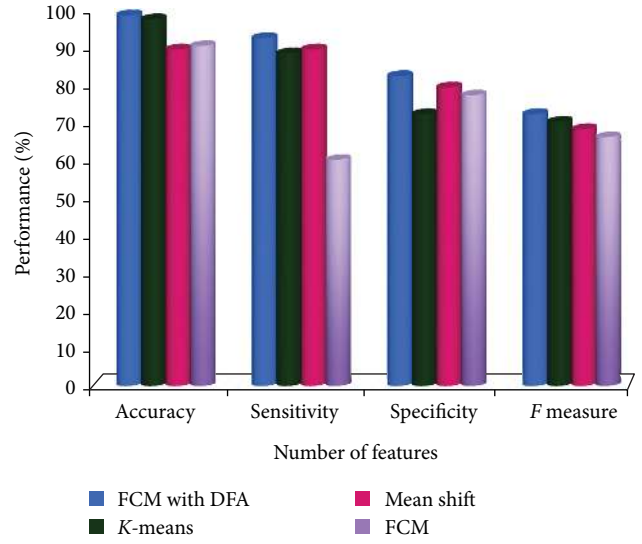

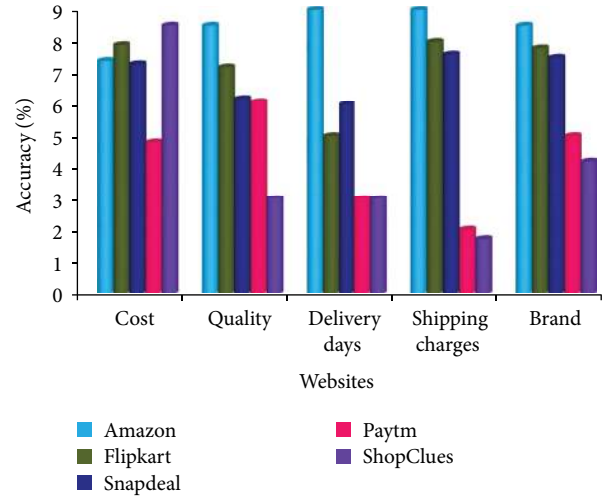FIGURE 4: Comparative analysis.

FIGURE 5: Ranking analysis.

however, the coveted ideal outcome is accomplished in the FCM with optimization.

With a specific end goal to assess the feasibility of sentiment sifting, we welcomed ten volunteers to physically rate the audits that have been sifted through. Figure 5 illustrates the changed class of item category, such as cost, quality, delivery days, shipping charges, and brand. Based on our analysis the best score was achieved by the Amazon site for most of the cases, so the vast majority of the clients like this site just for Internet shopping reasons. For the most part, the item having the most astounding rank may not really have great characteristics. This module incorporates the star evaluations, extremity of the reviews, age of the reviews, and the impact of item scores for calculating the score for an item.

7. Conclusion

Our approach shows and demonstrates the most effective web-based shopping sites and how they carried on. The

execution of an item in the wake of checking whether great or poor is determined by client surveys from business websites. Unhelpful surveys can be sifted through naturally from all buyer audits with a high review rate of about 88% and 90% accuracy. It is sincerely necessary for an organization to know the assessments of clients about its items. The positioning of items, item score, and the correlation of excess in between two items all prescribe an item list alongside its general score. In this paper we have thought about grouping features and advancement, like what we have used in positioning calculation. An organization must know which highlights of a specific item are required and which highlights should be enhanced to expand consumer loyalty. The main features on which opinion conveyed are selcted and the reviews are extracted based on the features identification. As a future improvement, the framework might reach out to utilize characterization procedures for positive and negative group surveys for better justification.

Data Availability

The data used to support the findings of this study are currently under embargo while the research findings are commercialized. Requests for data, 6 months after publication of this article, will be considered by the authors S.K. Lakshmanaprabu or K. Shankar.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] Y. Heng, Z. Gao, Y. Jiang, and X. Chen, "Exploring hidden factors behind online food shopping from Amazon reviews: a topic mining approach," *Journal of Retailing and Consumer Services*, vol. 42, pp. 161–168, 2018.

[2] S. A. Sadhana, L. SaiRamesh, S. Sabena, S. Ganapathy, and A. Kannan, "Mining target opinions from online reviews using semi-supervised word alignment model," in *2017 Second International Conference on Recent Trends and Challenges in Computational Models (ICRTCCM)*, pp. 196–200, Tindivanam, India, February 2017.

[3] S. H. Liao, Y. J. Chen, and Y. T. Lin, "Mining customer knowledge to implement online shopping and home delivery for hypermarkets," *Expert Systems with Applications*, vol. 38, no. 4, pp. 3982–3991, 2011.

[4] A. G. Babu, S. S. Kumari, and K. Kamakshaiah, "An experimental analysis of clustering sentiments for opinion mining," in *ICMLSC '17 Proceedings of the 2017 International Conference on Machine Learning and Soft Computing*, pp. 53–57, Ho Chi Minh City, Vietnam, January 2017.

[5] K. S. Kumar, J. Desai, and J. Majumdar, "Opinion mining and sentiment analysis on online customer review," in *016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, pp. 1–4, Chennai, India, December 2016.

[6] K. S. Oza and P. G. Naik, "Prediction of online lectures popularity: a text mining approach," *Procedia Computer Science*, vol. 92, pp. 468–474, 2016.

[7] S. S. Sohail, J. Siddiqui, and R. Ali, "Feature extraction and analysis of online reviews for the recommendation of books using opinion mining technique," *Perspectives in Science*, vol. 8, pp. 754–756, 2016.

[8] G. Li and F. Liu, "Application of a clustering method on sentiment analysis," *Journal of Information Science*, vol. 38, no. 2, pp. 127–139, 2012.

[9] X. Wang, X. Chai, C. H. Hsu, Y. Xiao, and Y. Li, "Cluster analysis based on opinion mining," in *2015 8th International Conference on Ubi-Media Computing (UMEDIA)*, pp. 110–115, Colombo, Sri Lanka, August 2015.

[10] G. Dubey, A. Rana, and N. K. Shukla, "User reviews data analysis using opinion mining on web," in *2015 International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE)*, pp. 603–612, Noida, India, February 2015.

[11] Z. Yan, X.-y. Jing, and W. Pedrycz, "Fusing and mining opinions for reputation generation," *Information Fusion*, vol. 36, pp. 172–184, 2016.

[12] S. S. Sohail, J. Siddiqui, and R. Ali, "Book recommendation system using opinion mining technique," in *2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 1609–1614, Mysore, India, August 2013.

[13] A. S. H. Basari, B. Hussin, I. G. P. Ananta, and J. Zeniarja, "Opinion mining of movie review using hybrid method of support vector machine and particle swarm optimization," *Procedia Engineering*, vol. 53, pp. 453–462, 2013.

[14] V. Singh and S. K. Dubey, "Opinion mining and analysis: a literature review," in *2014 5th International Conference - Confluence The Next Generation Information Technology Summit (Confluence)*, pp. 232–239, Noida, India, September 2014.

[15] N. Claypo and S. Jaiyen, "Opinion mining for thai restaurant reviews using K-means clustering and MRF feature selection," in *2015 7th International Conference on Knowledge and Smart Technology (KST)*, pp. 105–108, Chonburi, Thailand, January 2015.

[16] A. Parashar and E. Gupta, "ANN based ranking algorithm for products on E-commerce website," in *2017 Third International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB)*, pp. 362–366, Chennai, India, February 2017.

[17] S. Mirjalili, "Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems," *Neural Computing and Applications*, vol. 27, no. 4, pp. 1053–1073, 2016.

[18] S. Park and D. Y. Kim, "Assessing language discrepancies between travelers and online travel recommendation systems:

application of the Jaccard distance score to web data mining," *Technological Forecasting and Social Change*, vol. 123, pp. 381–388, 2017.

[19] M. Malika, S. Habiba, and P. Agarwal, "A novel approach to web-based review analysis using opinion mining," *Procedia Computer Science*, vol. 132, pp. 1202–1209, 2018.

[20] L. Xiao, F. P. Guo, and Q. B. Lu, "Mobile personalized service recommender model based on sentiment analysis and privacy concern," *Mobile Information Systems*, vol. 2018, Article ID 8071251, 13 pages, 2018.

[21] S. Aravindan and A. Ekbal, "Feature extraction and opinion mining in online product reviews," in *2014 International Conference on Information Technology*, pp. 94–99, Bhubaneswar, India, December 2014.

[22] Z. Zhai, B. Liu, H. Xu, and P. Jia, "Clustering product features for opinion mining," in *WSDM'11 Proceedings of the fourth ACM international conference on Web search and data mining*, pp. 347–354, Hong Kong, February 2011.

[23] Z. Cebeci and F. Yildiz, "Comparison of K-means and fuzzy C-means algorithms on different cluster structures," *Journal of Agricultural Informatics*, vol. 6, no. 3, pp. 13–23, 2015.

[24] S. R. KS and S. Murugan, "Memory based hybrid dragonfly algorithm for numerical optimization problems," *Expert Systems with Applications*, vol. 83, pp. 63–78, 2017.