Taylor & Francis
Taylor & Francis Group

Automatika
Journal for Control, Measurement, Electronics, Computing and Communications

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/taut20

A scheme of opinion search & relevant product recommendation in social networks using stacked DenseNet121 classifier approach

Murugesan Shanmugavelu & Muthurajkumar Sannasy

AUTOMATIKA
2023, VOL. 64, NO. 2, 248–258
https://doi.org/10.1080/00051144.2022.2140389

REGULAR PAPER

OPEN ACCESS Check for updates

A scheme of opinion search & relevant product recommendation in social networks using stacked DenseNet121 classifier approach

Murugesan Shanmugavelu[a] and Muthurajkumar Sannasy[b]

[a]Department of Information Technology Tagore Engineering College, Rathinamangalam, Chennai, India; [b]Department of Computer Technology, Anna University, Chennai, India

ABSTRACT
Traditional methods have resulted in lower-quality search results with a lower accuracy rate. This problem is addressed and a recommended technique using deep learning methods is provided with the goal of improving prediction quality. Via this paper, a novel paradigm for pertinent product recommendations in social networks is provided. The major goal of this strategy is to let computers learn automatically without any assistance from humans, consequently controlling operations as needed. The social input data set in this proposed study is first pre-processed to remove noise. Following that, a Fisher discriminant method based on information is used for feature extraction. Then, using the Hierarchical Agglomerative and Attribute-based Clustering procedure, the features are chosen from the retrieved ones. Following that, such clusters are predicted using the stacked DenseNet121 method, and Attention-based MLP is used to propose the product. Finally, to verify the effectiveness of the suggested system, the expected output was evaluated, the performance measure was examined, and comparisons with conventional methods were made. Out of 2033 reviews, the suggested approach has a positive score percentage of 92.22%. The investigation demonstrates that the suggested system is more effective at providing improved results for pertinent product recommendations.

ARTICLE HISTORY
Received 28 January 2022
Accepted 20 October 2022

KEYWORDS
Recommender system; relevant product recommendation; information based fisher discriminant algorithm; hierarchical agglomerative and attribute-based clustering; stacked DenseNet121; attention-based MLP

Introduction

The amount of data has steadily increased over the past few years. Companies began to keep a wide variety of data, including server logs and any other helpful information with commercial worth. As a result, we now have a significantly greater capacity to gather data from numerous apps in various formats. Big Data refers to these data and the systems that can handle them, and it has significantly impacted today's company. A few years ago, a company would keep a restricted subset of data that only contained essential facts. In contrast, a business may now use huge data to harness all of the information at their disposal in order to acquire insights and improve decisions, giving them an edge over rivals in the market. One of the best examples of how Big Data is used in everyday life is recommender systems. Services employ recommender systems approaches to mine and process vast volumes of data for applications like e-commerce and music/video streaming to better fit the demands of its users in a personalized way. Recommender systems have primarily developed as a technique to assist users in their decision-making process since they recommend the most appropriate goods to a specific user. The term "recommender system" (RS) refers to a group of tailored algorithms that determine each user's preferences using machine learning (ML) and data mining techniques. Recommender systems are frequently used in e-commerce as a differentiator.

In order to provide meaningful product recommendations in the social network, a deep learning-based system was provided. Traditional methods have resulted in lower-quality search results with a lower accuracy rate. This problem is addressed and a recommended technique using deep learning methods is provided with the goal of improving prediction quality. The main goal was to make it possible for computers to learn automatically without the assistance or instruction of humans, hence controlling the operations. The social input data set was first pre-processed in this suggested effort to remove noise. Fisher discriminant algorithm based on information was used for feature extraction. Then, a hierarchical agglomerative and attribute-based clustering technique was used to choose the features. This HAC is easier to deploy, and the development of a dendrogram is beneficial to understanding how the clusters are now organized hierarchically. Any such legitimate distance measurements that are less influenced by cluster morphologies that are sensitive to handling are approved. Clusters with different sensitivities are capable of being visualized well. Then, using stacked DenseNet121 classifiers, a prediction was created, and a product recommendation was

CONTACT Murugesan Shanmugavelu muruga13@gmail.com Assistant Professor Department of Information Technology Tagore Engineering College, Rathinamangalam, Chennai, India

made using attention-based MLP. In order to verify the effectiveness of the suggested system, the expected output was evaluated and the performance estimation was compared with conventional methodologies. The development of online shopping is increasing in the recent world. The main feature of online shopping is shopping the ease of the house. Recommender systems RS have been employed by different domains such as e-commerce, social websites, etc. On e-commerce websites like Amazon, Flipkart, and Yelp, users can share their feedback and opinions about the product purchased. Products are recommended according to their personal interest or similar to be bought in the forthcoming days. The suggestions depend on the customer opinion of who bought the item and the opinions were stated in the form of ratings or reviews thus influencing the customers. The consumer's interest was seized by the list of features that belongs to the methods such as content-based filtering, metadata-based filtering, collaborative filtering (CF), etc. Reviews are much significant for businesses as they increase sales by offering suggestions to buy the item and it plays a vital role in increasing the standard, and reputation of an e-commerce store. Product reviews are required for both the customers and sellers [1]. Many manufacturers and new items were received online daily which creates conflicts among the users while buying the product. Customers consider the former user reviews to make a purchase. The probability of a Product being sold is high if the ratings are high for the specified item.

The main concept is to generate a recommendation system that is capable of recommending the relevant item. Reading all the reviews and comments consumes more time than we call information overload. Text summarization is the method of summarizing all the reviews into a summary and is divided into Extractive and Abstractive. Abstractive summarization reviews the file on its own and summarizes the events. In Extractive summarization, sentences are rated and a summary is made by using the top-rated sentences [2]. The recommending system offer content based on their needs to the customers by searching a large amount of data and also enhances the ability of decision-making. It helped our consumers to identify the record based on their preferences. Content-based filtering is an algorithm that highlights the attributes of the resources for producing the best resources from the list of records. The recommendation system altered the method of communication between websites and consumers and suggests the appropriate product to the customers [3]. The recommender system cannot be neglected for e-commerce websites and advertisements. They have large applications in various fields like economy, education, and research. This system has achieved the best result by solving the issue of information overload. Since there are many products on e-commerce websites, it is difficult for the customers to select their specified choice. The recommender system could predict whether the customer would select the product or not on the basis of the customer's profile. The process of making decisions and quality has been enhanced through Recommender systems that result in the effects such as identifying the product according to the user's interest, benefits the seller to distribute the product to the correct customer, finding the relevant product, and supports the website to enhance the engagement of the customer [4]. The limitations of online shopping are the touch and feel parameters of direct shopping. The user requires some factors like customer reviews and feedback to decide about the item. The user cannot buy the product by seeing the images. A product with more positive reviews may lead to an increase in sales. Many recommendations system consider the reviews to offer consistent information for the consumers whether to purchase the item or not [5]. E-commerce websites like Amazon and eBay have social network features consisting of real-time updates and communications among the sellers and the consumers. These websites also permit new customers to log in with their credentials from social networking like Twitter, Google, and Facebook. We suggest knowing the customer characteristics illustrations (Utilizer embedding) and product characteristics illustrations (Product embedding) from the collection of e-commerce websites that uses the recurrent neural network and then the gradient boosting trees technique is employed to alter the customer's gregarious networking characteristics into utilizer embedding. A feature-based matrix factorization method is implemented to control the utilizer embedding for cold-start recommendation. The outputs on the dataset have displayed the efficiency of the proposed work in which the recurrent neural network is applied to understanding the characteristic representation of the item and the customers [6]. Abundant fluctuations in the way of our lives have been brought about by the internet or the World Wide Web. There is a tremendous rise in the volume of digital information and the users of the internet have made probable challenges of overload in the information on the consumers trying to identify the relevant product. Hence the implementation of the Recommender system has been initiated that can suggest whether a customer would desire the product. RS is a process that can recommend the most relevant item to the specified consumers by understanding the interest of the customer in the product on the basis of relevant information regarding the product. The focus of the RS is to reduce the overloaded information by extracting the most relevant items from the large volume of data and offering custom-made services to the customers. The significant characteristic of the RS is to guess the interest and preferences by understanding the activities of that customer to produce recommendations [7].

The left-over section of the paper is arranged as follows: section II is the study of various already existing techniques and reviews used. Section III is about the proposed work. Section IV shows the performance analysis of the proposed method and the comparative analysis of the outcomes. At last, section V provides the conclusion.

Related works

Shaikh et al. [8] described the scarcity of semantic factors in the system of recommendation and the various methods of recommendation that are being used on the websites of e-commerce. Systems of content based consider the item properties to be suggested. For example, if a user of Amazon had bought many comic books, then the database having the "comic" genre was recommended by the content-based system. Collaborative systems suggested the product on the basis of measures of resemblance among like-minded customers. This paper emphasized the requirement of semantics to suggest the items and described different limits in the recent recommendation methods. This also offered solutions to enhance the system used on the websites of e-commerce. It included many websites such as eBay, Amazon, Flipkart, and Paytm by giving a rating on the basis of various parameters and comparative analysis is done. The suggested system compared the recommendations of flickr.com images with the proposed technique. This implemented the semantic recommendation based on the graph with the help of the overlap method.

Singla et al. [9] mentioned the reviews that play a vital role to develop communications among the users and thereby influencing buying patterns. Websites such as Amazon, and Flipkart offer a platform for users to give an opinion about the efficiency of the product to future consumers. Classification of reviews into negative and positive sentiment is needed to get insights from huge reviews. Using sentiment analysis, over 4 lakh reviews were classified into negative and positive sentiments. Support Vector Machine, Naïve Bayes, and decision tree were the classification models used for reviews, and computation of models is performed by 10-Fold cross-validation.

The aim of this paper is to identify the same kind of users and provide a suggestion about a product based on the rating [10]. The customer could buy a specified product if it is available online. The choice of selecting an item is huge since many items where many products are available. Sometimes the consumer did not know the persistence of the relevant product. In order to solve the problem websites, use many various strategies to offer a recommendation to the consumer. Products were recommended based on the customer's personal interest. Here, a verification step was implemented to place the reviews for the items purchased online. Only the customers who have bought the item could post their opinions and feedback.

Deng et al. [11] introduced an algorithm of recommendation on the basis of the diffusion method that integrates the social network and customer item relations. Recommender systems and social networks were the most internet applications and they were under a single framework because of the nature difference. The information in the social networks was related to the recommender system. The results showed that the improvement in the accuracy of recommendations was done by social networks and the variety of recommended items for the active customers increased. It performed algorithms based on conventional popularity and recommendations are provided during the cold start period. These outputs highlighted the algorithms for recommendation in integrating the recommendations and social networks.

Kiran et al. [12] proposed an alternative for product rating on the basis of its technical specification by considering huge customer reviews that are extracted from the various e-commerce websites. Product recommendation plays an important role to attract customers based on their requirements. There was no need to search for comments and opinions, before purchasing a product. In this approach, the analysis is done to extract the list of specifications such as processor, battery, camera, and user reviews for a particular item from various websites. This also identified the critical terms related to the technical features of the item to regulate the polarity. Based on polarity, each and every specification is assigned positive or negative feedback. Overall item rating is estimated by accumulating the score specific. This method was useful for users who require a particular feature in the product.

Dadhich and Thankachan [13] addressed and reviewed the concepts of identifying the sentiments automatically stated in the products of Amazon, and Flipkart using Random Forest, Naive Bayes, Logistic Regression, K-Nearest Neighbour techniques, and SentiWordNet. The retail market has engaged to sell the items online and also offer their suggestions, recommendations, and feedback. The system of classification and the summarization of opinions and feedback recognize the opinions about the various online products in a list of text-based reviews. This presented a comparative analysis of already existing sentiment algorithms based on the key parameters. This paper suggested the Product Comment Summarizer and Analyzer system (PCSA) which is an analyzer of generic comments that can identify the sentiment polarity in an efficient manner. The comments were summarized and classified into the pre-defined negative, positive or neutral and the results are evaluated according to efficiency based on the parameters such as classifiers, rating, and accuracy.

Pothuraju et al. [14] adopted the architecture suggested in Caser [1] and the embedding of two products was introduced, new convolution blocks have been established, and a 3D tensor is formed with the embeddings. Recommender systems are employed to suggest the items to the buyers according to their preferences. To design the short-term behaviour of the customers, a sequential recommender is employed. Many methods like the Markov chain and CNN convolutional neural networks have been employed to solve the problems of sequential recommendation. The top-N sequential recommendation system is the most effective state-of-the-art model using CNN. This method failed to seize the skipping behaviour since it did not contain the distance items interaction.

This approach increased the MAP mean average precision that gives rise to the sequential recommendation system.

Mongia [15] provided statistical insights into reviews of the items on Amazon, analyzed their utility in product suggestions, and employed a new technique for the prediction of the utility score of the reviews of the customer. Processing abilities have trapped to sue the large quantity of data with many algorithms.

These methods were worthy for retailers online, using the rating to offer a reputation in e-commerce. Quantitative ratings and text reviews were both provided by many online shops. Reviews were increased over the years on websites of e-commerce. Purchased Users on Amazon rated the items up to 5 stars and the summary is shared in the form of text by sharing the opinion, feedback, and experience about the item.

DWIVEDI [16] employed recommendation systems that are built using collaborative filtering and popularity-based system. The choices of selecting products are increased with the increase in e-commerce. Customers needed the recommendation system to identify their favourite and the best product from large resources. It was essential to discover the products by the customer thereby increasing the company sales. Models are computed using MAE and RMSE.

Pote et al. [17] explored the sentiment analysis, and used the influence of sentiment analysis that benefits both the seller and the user. E-commerce sites such as Amazon and Flipkart are visited by the users for purchasing the item. Reviews and genuine comments were important for businesses, sellers, and customers. Sellers should know what customers feel about the products and services. As there were a lot of comments about an item, it is not feasible to read each and every post and get the viewpoint.

Processing of large data was allowed by the sentiment analysis effectively and was the cost-efficient method to study the sentiments. Reviews of the users who already purchased the product were viewed by the buyers and also analyzed by the seller to develop their item to sell better.

Paranjape et al. [18] encountered recommender systems in our life during the communication with many online services. The recommender system plays a vital role to suggest a specific item to the buyers based on their interests. Various methods were implemented in e-commerce-based systems like Amazon, Flipkart, etc. Here, the suggested work deployed an approach of machine learning for recommending the product based on the collaborative method. The concept of SVD (Singular Value Decomposition) was developed to receive ratings and to reduce the issue of scarcity. In this paper, an evaluation of Root Mean Square Error (RMSE) is performed to verify the accuracy of the suggested work.

Camacho and Alves-Souza [19] showed the outputs of a review on Systematic Literature on Recommender System based on Collaborative Filtering which used the data of social networks to alleviate the issue of cold start. The most relevant item suggested that meets the requirement of the customer was achieved by the Recommender System that deals with the overload due to the huge amount of data. The system could not recommend the products when the customer was new to the system due to the lack of previous interest of the consumers and the history of rating to determine the preferences. The above issue is the cold start problem which did not have a solution. To solve the cold start problem social networks have been implemented that provides a source of information to regulate preferences. The final outcomes proved that many publications use social network information. Few papers use this data to alleviate the problem of cold start.

Portugal et al. [20] analyzed the use of algorithms based on machine learning in the recommender system and found the chances of study for research of software engineering. The recommender system used machine learning algorithms from AI (Artificial Intelligence) to offer the consumers the recommendation of the item. As there were many algorithms, it was hard to select a particular machine learning algorithm. The implementation of the Recommender system using a machine learning algorithm has many issues, and hence the engineers recognize the aim of the efforts. This concluded that decision tree and Bayesian algorithms were employed in the systems due to their relative simplicity and the design phases and requirement stages offered the ways for future study.

The extraction of artificial features also has an impact on classification accuracy. The relative length of the online text makes it more likely that the classifier will miss critical information, which reduces the model's capacity for detection Jing-Yu and Ya-Jun [21]. In recent years, human interests have shifted from the land to the sky and the sea. Robots could be used for dangerous jobs that humans are unlikely to fulfil, including sending humans into space and the ocean to explore. Recently, humankind has turned its attention from land-based research to cosmic and oceanic exploration. Activities that put humanity in danger, including sea and space research, are done by robots [22].

We have reached the following conclusions using the most recent state-of-the-art friend recommendations. Some existing efforts that use "trust" are based on trust-propagation or users' trust-based neighbours to identify user communities. However, It's important to note that user trust information is rarely available in social networks. This is why we believe formalizing is more important.

The idea of trust for a certain person in a social network taking into consideration his/her interactions with other members. According to our knowledge, no research has yet addressed how a user's network credibility measures how committed they are.(i.e. involvement and sociability) as well as his or her level of trustworthiness based on their expertise, value, and seniority. Typically, a user's "tastes" (i.e. similarity amongst users) are determined by the ratings or tags they give goods on a website extracted themes or even lifestyles of users. On the other hand, past study on friend recommendations primarily focused on connecting relation, social media users, but mostly ignored the impact of the individuals' characteristics. In actuality, we have observed that only few. For the purpose of user recommendation; works have merged semantic and social network relations. Several authors combined social friendships to provide recommendations without differentiating between them. Between various friendships between users (such close/immediate or distant pals). Lastly, several efforts have added significant elements like the classification in addition to the concepts of trust, similarity, and social links and user clustering, context, and influence (in its various guises, such as local and remote influence, which allude to influence from close friends and afar, respectively).

We shall emphasize the following elements in this paper. The implicit profile of a user is modelled through social network interactions. In the context of social networks, the suggestion of possible friends to a specific user makes use of the semantic and social aspects of individuals' profiles while taking their trustworthiness into account (their degree of trust and commitment). Our contribution will be set apart from other already-in-use systems by the incorporation of users' credibility, which uses both the social and semantic components of the recommendation. Grouping comparable people based on semantic and social attributes using classification techniques to suggest close and related buddies. We think that combining the social and semantic components will improve accuracy. Traditional methods have resulted in lower-quality search results with a lower accuracy rate. This problem is addressed and a recommended technique using deep learning methods is provided with the goal of improving prediction quality. Via this paper, a novel paradigm for pertinent product recommendations in social networks is provided. The major goal of this strategy is to let computers learn automatically without any assistance from humans, consequently controlling operations as needed.

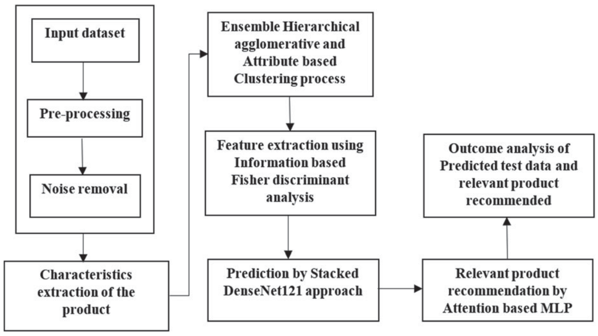

Figure 1. Entire workflow of suggested scheme.

Proposed work

A brief description of the proposed workflow mechanism is narrated in this section. The entire workflow of the suggested scheme is shown below in Figure 1.

Input data pre-processing

A process of data cleaning is carried out in the pre-processing step so as to make the data noise-free. The data noise is in the form of unwanted or missing words. Those unwanted or missing words are in the disguise of a few symbols that are not handed over by code. This data cleaning consists of filtering of information, tokenization process, and removal of punctuation or stopwords. Once the cleaning process is over, the data is then converted to lowercase, stopwords are removed, and tokenized. Figure 2 shows pre-processing of data comprises parts of speech, tokenization, normalization, stemming of words, and lemmatization at which the network among the words is established. By this, the sub-events and relevant events are entirely mapped into the connecting words for instance: thanking you, thank you could be mapped into thank you. The clean data is subsequently subjected to this stemming procedure. Pre-processing is used to lower the data's noise level. Thus, opinion mining includes the ensuing

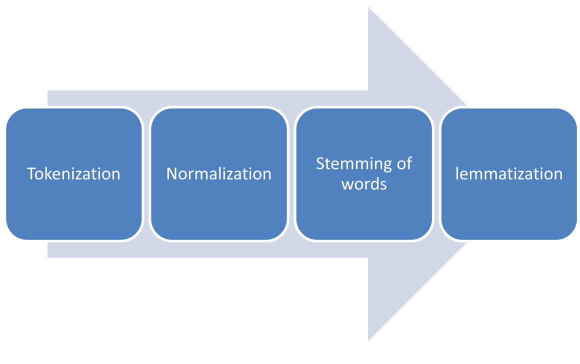

Figure 2. Flow chart of preprocessing techniques in speech data.

pre-treatment processes, including tokenization, word removal, stemming, cleaning, and so forth. This pre-processing step is employed for the purpose of reducing the number of characteristics or features which aids in obtaining the selected method in a successful manner. A primary step of this pre-processing is the tokenization process at which the input data is spitted as small pieces of words termed tokens which eliminates the low priority aspects. Then, the word deletion is stopped which aids in decreasing the matrix size and thus augments the degree of discrimination between the words. Then, the stemming process is employed which translates single stems to separate ones for deciding whether it is text or word stems. It is the process of dropping a word stem that sticks to the suffixes and prefixes or the roots of words recognized as the lemma. The sentence is segmented with respect to the stop word. The stop word is a term that is used frequently when the search engine has been automated, both when the access index for incisive and the one that revives them as the outcome of the search query.

Ensemble hierarchical agglomerative and attribute-based clustering process

After preprocessing, the characteristic of words is extracted and the clustering process is done using Ensemble Hierarchical Agglomerative clustering (HAC) and an Attribute-based clustering scheme. The HAC is regarded as an unsupervised algorithm that works in a "bottom-up" manner. This HAC technique is an easy and simple means to understand and there is no such necessity to stipulate the number of the cluster for this process. The implementation of this HAC is a simpler one and for its interpretation, the dendrogram production is helpful and beneficial to, the current relation of hierarchy among the cluster. Any such valid measures of distance are accepted that are less influenced by means of cluster shapes that are sensitive less for handling. Clusters having various sensitivity are having well visualization capacity. The algorithm steps of this HAC are as shown:

Step 1: At first, each data points make a single piece of cluster elements or leaf.

Step 2: Estimate the matrix distance among clusters.

Repeat

- A two nearest or closest ones are united at each point of algorithm
- The distance matrix is then modernized

Until

- The entire data points are just a part of remaining single cluster that is the root cluster

Step 3: The single link distance is then estimated as: The distance among clusters kj and ki must be a least distance among any such objects xi and xj.

$$Sim\,(kj, ki) = mini\ sim(xj, xi) \tag{1}$$

where, $xj \in kj, xi \in k$.

This could outcome in long because of effect of chaining and carried in o(n)^2. Here, n signifies the datapoints number.

Attribute-based clustering is a process that recognizes c clusters disjoints like $C_1, \ldots C_c$ of the attributes that are correlated through assigning every attribute in the $A_1, \ldots, A_p$ for the one of those clusters. Typically, attribute-based clustering is a process that is defined as that $\forall \mathrm{A}_i,\ i \in \{1, \ldots, p\},\ \mathrm{A}_i$ is being assigned to the $C_r, r\epsilon\{1, \ldots, c\}$ at which $C_r \cap C_s = \phi$ for entire $s \in \{1, \ldots, c\} - \{r\}$.

So as to recognize the meaningful clusters, attribute-based clustering is also performed such that the attributes in the cluster must have higher correlation or having higher independence to one another while the attributes of various clusters are correlated less and or highly independent.

In attribute-based clustering, each tuple in relation R is categorized through attributes set like $A_1, \ldots, A_p$. In the case of $\mathrm{A_i}, i \in \{1, \ldots, p\}$ takes discrete values, let their domain be signified by $dom(A_i) = \{a_{i1}, \ldots, a_{imi}\}$. Or else, in case $\mathrm{A_i}, i \in \{1, \ldots, p\}$ is continuous, let their domain be signified by means of $dom(A_i) = [l_i, u_i]$

Here, $l_i, u_i \in \Re$ consider that R comprises of n number of tuples like $t_1, \ldots, t_n$. At each tuple, $t_u, u\epsilon\{1, \ldots, n\}$ is then signified through the p attributes vector values: $t_u = (x_{u1}, \ldots, x_{up})$ at which$x_{ui} \in dom(A_i), i = 1, \ldots, p$.

Thus, the integration of these two contributes the ensemble HAC and attribute-based clustering.

Information-based Fisher discriminant analysis for feature extraction process

In this segment, Information based Fisher discriminant analysis is implemented for extracting the features. Extraction of features is performed from the information of the preprocessed data. Information-based Fisher discriminant analysis is regulated and managed method and the information is utilized for extracting the features indirectly.

Suppose the number of matrix patterns of classes is N.

$$\omega = \{A_{pq}\}_{q=1}^{T_p}, \quad p = 1, 2, \ldots, N \tag{2}$$

which denotes the pth class. The class mean is

$$\overline{A_p} = \frac{1}{T_p} \sum_{q=1}^{T_p} A_{pq}, \quad p = 1, 2, \ldots, N, \tag{3}$$

A denotes the sum of the mean samples.

Let us assume a vector m. Information based Fisher Discriminant analysis estimates A, a pattern onto the vector m satisfying the constraint $m^T m = 1$ with the help of linear transformation

$$s = m^T A$$

s is an extracted feature matrix. So each A_{pq}, $p = 1, 2, \ldots, N; q = 1, 2, \ldots, T_p$, all the predicted values are in the type of

$$s_{pq} = m^T A_{pq}, \quad p = 1, 2, \ldots, N; \quad q = 1, 2, \ldots, T_p \tag{4}$$

To get optimal vector m, we denote the real function as

$$R_{Mat}(m) = \frac{tr(m^T E_b^{Mat} m)}{tr(m^T E_w^{Mat} m)} \tag{5}$$

Where

$$E_b^{Mat} = \sum_{p=1}^{N} T_p(\overline{A_i} - \bar{A})(\overline{A_i} - \bar{A})^{\mathrm{T}} \tag{6}$$

is the total of between classes of scatter matrix and

$$E_w^{Mat} = \sum_{p=1}^{N} \sum_{q=1}^{T_p} (A_{pq} - \bar{A}_p)(A_{pq} - \bar{A}_p)^{\mathrm{T}} \tag{7}$$

the total within classes of scatter matrix.

We attain two points in the projection space by increasing $R_{Mat}(m)$. One point is to create between scatter class as big, other point is within scatter class as small.

Differentiating $R_{Mat}(m)$ with respect to m under the condition of $m^T m = 1$, the following equation of eigenvalue-eigenvector that m satisfies

$$E_b^{Mat} m = \lambda E_w^{Mat} m \tag{8}$$

Prediction using stacked DenseNet121 approach

In stacked DenseNet121 all the layers are connected directly to ensure the maximum rate of transmission of information between the network layers. The advantages of DenseNet121 are a narrow network and fewer parameters. Each layer receives additional input information from the former layers and transmits it to all the successive layers so as to maintain the characteristics of forwarding propagation. In DenseNet121, each adjacent layer is connected with the former layer directly and hence there is P(P+1)/2 connections. DenseNet121 consists of multiple Dense Blocks that are linked by a transition layer which in turn is composed of a BN layer with a 2×2 pooling layer and a 1×1 convolution layer.

Let us assume the size of the input data as M x M x C. After it is transferred from the bottleneck layer, then ReLU and 3×3 convolution layers the output size is M x M x 4.

The input of the second layer depends on both the input and output of the first layer and the size is M x M x (C+4). Then it is transferred and processed by BN, ReLU, and 3×3 convolution layers. Then the previous function is repeated and the output of the fifth layer is M x M x 4. The input-output relationship can be denoted as

$$x_n = F_n([x_0, x_1, \ldots\ldots\ldots x_{n-1}]) \tag{9}$$

$[x_0, x_1, \ldots\ldots\ldots x_{n-1}]$ is the previous layer feature splicing. F_n is the composite function of 3 successive continuous operations.

The stacked densenet121 classifier algorithm is provided below:

Algorithm 1: Stacked DenseNet121 Classifier algorithm for prediction purpose

---

Input: features selected with current weights w_{ij}^m, *training samples*
Output: New weight and classified output
Initialize the following steps
1. Each samples n are computed with internal and output activation units
2. Calculate individual unit x_i^m *as the term error propagation* $\beta_i^m(n)$ *envisioned for the output layer as follows:*
$\beta_i^{l+1}(n) = |(c_i(n) - x_i(m))|_v = z_i^{l+1},$
At which, $m = l+1$
3. Evaluate the term error propagation $\beta_i^m(n)$ *destined for the hidden layers as exposed:*
$$\beta_i^m(n) = \left| \sum_{i=1}^{N^{m+1}} \beta_i^{m+1} w_{ij}^m \frac{\partial fv}{fv} \right|_v = z_j^m$$
Here, $v = z_j^m = \sum_{j=1}^{N^{m-1}} y_j^{m-1}(n)\, w_{ij}^{m-1}$
4. the concern of weight connection is specified by the subsequent equation as follows:
$$\text{new } w_{ij}^{m-1} = w_{ij}^{m-1} + \delta \sum_{t=1}^{T} \beta_i^m(n) \cdot y^{m-1} \cdot j(n)$$
end

Thus, the classification based on reviews is made by use of this stacked densenet121 classifier approach. Then the relevant product recommendation is done with the use of an attention-based MLP classifier mechanism.

Relevant product recommendation by attention-based MLP

MLP is typically a feed-forward neural network that is trained using a standard back-propagation method. This method is frequently used for problems with prediction modelling and allows training to learn the input transform data as a preferred response. The Attention-based MLP model, which comprises a number of characteristics and functional layers, is used in this work to improve the MLP by establishing the categorization outcomes that are significant. The functionality of this attention-based MLP is classified as focused reviews and appropriate products for suggestion and is based on kernel function. Any linear model at each enabled kernel function is twisted by applying a kernel trick to the non-linear template. All those attributes are then determined by the thresholds of the kernel such that the higher thresholds and attributes are vastly provided. This approach is selected since it is having better generalization abilities and is stable extremely that works better for data sets. This can be expressed as follows:

$$\mathrm{T}_{\mathrm{j}}^{\mathrm{m}} = \sigma \sum_{\mathrm{L}} \mathrm{x}_{\mathrm{jk}}^{\mathrm{m}} \mathrm{a}_{\mathrm{K}}^{\mathrm{m}} + \mathrm{b}_{\mathrm{j}}^{\mathrm{m}} \tag{10}$$

The vectorized equation is shown as follows:

$$\mathrm{b}_{\mathrm{j}}^{\mathrm{m}} = \sigma(x^m a^{m-1} + b^m) \tag{11}$$

The set of quadratic for combining training set is exposed as,

$$q = \frac{1}{2}||Y - a^m||^2 = \frac{1}{2}\sum_j (y_j - a_j{}^m)^2 \tag{12}$$

The output of gradient is provided as,

$$\frac{\partial \mathrm{C}}{\partial \mathrm{w}_{\mathrm{Lj}}^{\mathrm{m}}} = \mathrm{aL}^{\mathrm{m}-1}\delta_{\mathrm{j}=\mathrm{C}}^{\mathrm{m}} \tag{13}$$

Therefore, the transformation is expressed as,

$$Transfrm = \sum_{fea\,(x,y)} \left[W_{match(1)}x - \frac{fea}{2} \Big/ match + W_{fea(2)}x - fea/_{match} \right] /text$$
$$\mathrm{number(n)} * \frac{1}{\sqrt{2 * \pi * \delta_{ev}{}^2 * exp(\psi_{em} - 1)^2 \delta_{ev}{}^2}} \tag{14}$$

Once the query has been provided, then the several sub-data for the entire query in the dataset is being formed. The dataset is then numbered as 1, 2 ... and so on to right. After that, the query data design of a similar name in the archive and request is expressed as follows:

$$objctve_{ND} = \text{features transformed} + \text{score value} \tag{15}$$

Here, ND denotes the matching distance of Euclidean, q signifies the query data and t denotes the score value of data. Based on the score value, the relevant products according to the search and reviews are predicted and recommended.

Performance analysis

Performance analysis of the proposed system is estimated and the outcomes are projected in this section.

(A) Dataset Description
- Facebook – Nearly 1000 posts on Facebook of Amazon for “Kindle” products are collected
- Twitter – 4000 reviews on Kindle products collected
- Amazon – Around 1.5 million reviews are available and 4000 reviews are considered

(B) Performance and comparative analysis

The suggested system executes on the Amazon dataset, Tweets, and Facebook that has contained nearly 10,000 reviews on Kindle products. The expression including likes, satisfaction, dislike, interest, and feedback are considered attributes in the database. Table 1 represents the dataset attributes.

The users convey their opinion on the products purchased online through certain information on social networks. On Twitter, Facebook, and Amazon, slang words, and emojis are used by customers to indicate their satisfaction or dissatisfaction with the items. Thus, the attributes are the expression that the people give comments about products on social networks.

On the basis of Amazon reviews, Tweets, and Facebook posts of “Kindle” products, the result is tabulated in Table 2 the opinion, tweets, and posts without sensational words are left out by investigation. The preprocessing reviews contain the opinion of sensational words.

The graphical representation of reviews on Amazon, Tweets, and Facebook according to the volume, preprocessed reviews, and positive and negative reviews are given in Figure 3 This stated that the product from Amazon is the most recommended by the users to buy.

The integration of the Amazon Reviews, Tweets, and Facebook Comments are done and the experimental results are shown in Table 3. The method of integration is performed and verified on the separate volume of reviews about the product. The investigation result showed the entire opinion of users on the item.

Table 1. Dataset attributes.

S. no.	Attributes	Count
1	Negation	104
2	Intensifier	Based on reviews
3	Conjunction	4
4	Emoji	751
5	Slang word	290
6	ReTweet	Based on Tweets
7	Modifier	92
8	AFINN	18,543
9	Likes	Based on Tweets
10	Overall rating	Based on Amazon reviews
11	Facebook’s text review	1000
12	Tweet’s text review	13,000
13	Amazon’s text review	13,000

Table 2. Study results based on Amazon, Tweets, and Facebook Reviews.

S. no	Sources	Amazon	Tweets	Facebook
1	Preprocessed reviews	2289	3294	747
2	Volume	3000	4000	1000
3	Positive reviews	2253	1777	565
4	% of positive reviews	98.43	53.95	75.64
5	Negative reviews	36	1515	182
6	% Of negative reviews	1.57	45.99	24.36

Table 3. Study results based on customer comments of Amazon, Tweets, and Facebook.

S. no	Amazon reviews	Tweets	Facebook reviews	Total reviews	Positive reviews	% of positive reviews	Negative reviews	% of negative reviews
1	3000	4000	1000	6369	4765	74.82	1604	25.18
2	4000	4000	1000	7034	5274	74.98	1760	25.02
3	1000	4000	1000	4967	4013	80.80	954	22.75
4	2000	4000	1000	5714	4401	77.02	1313	22.98

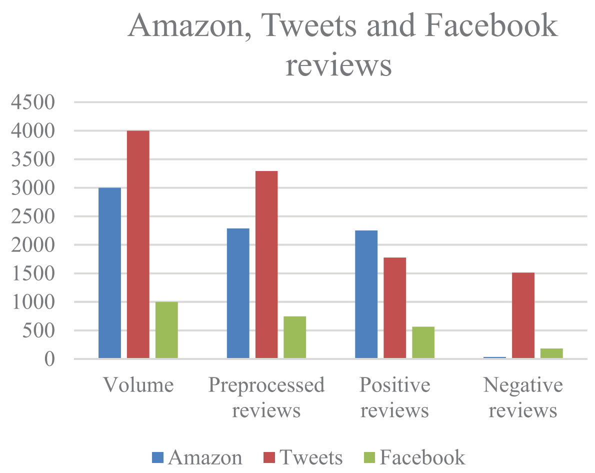

Figure 3. Reviews of Amazon, Tweets, and Facebook.

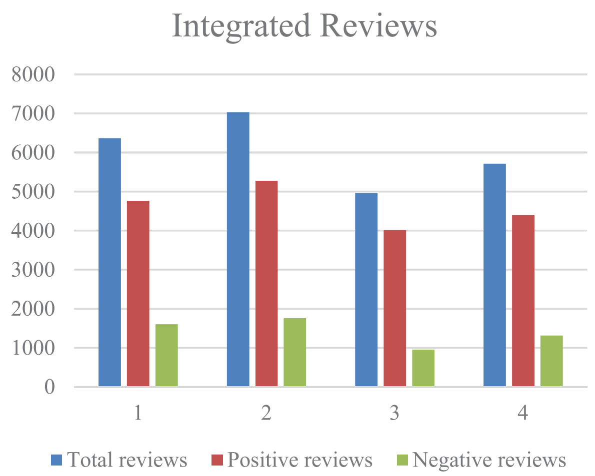

Figure 4. Integrated reviews on the product.

Figure 4 illustrates the reviews on total, positive and negative (integrated reviews) on the product. This stated that the recommendation for the product by the consumers is high.

These tests are primarily categorization. Particularly binary categorization where the result is always boolean, meaning that it can only be True or False. Simply because the result is True or False does not imply that the classification is 100% accurate. Almost all medical tests include a margin of error and are not always correct. One can anticipate getting very close to 100%. And the result is determined by how close it is to being true or untrue. As a result, when something is classified using binary classification, one of the four categories is selected as the output. True Positive (TP), True Negative (TN), False Positive (FP), and False Negative are listed in that order (FN)

Table 4. Comparison of performance.

Methods	Precision	F1-score	Accuracy
Stacked DenseNet121 + attention-based MLP (proposed)	100	97.65	98.45
Decision Support System [13]	100	94.04	91.57
SVM [5]	91.65	95.53	91.85
DT [7]	92.88	91.88	85.98
NB [8]	89.32	94.36	89.57
Rule Based Sentiment Analysis [17]	100	95.92	93.28

Precision: Out of all positives, how many were accurately categorized as positive.

Precision = TP/TP + FP

F1-Score: The F1 score is thought to provide a more accurate representation of the classifier's performance than the standard accuracy metric.

F1 score = 2 ∗ (precision ∗ recall)/ (precision + recall)

Accuracy: The percentage of successfully classifying the samples to their classes is determined by the model's accuracy. As it depends on the data and class balance, it cannot be used as a particularly good measure for verifying a model. We cannot say the model is functioning effectively if a given class is a minority and accuracy is 99 percent, which is primarily achieved by forecasting the majority class. However, the disparity in class is a totally different topic, and we shall talk about it later. Although it is not a circumstance that occurs in real life, the accuracy model is preferable to apply when there is no class imbalance.

Accuracy = (TP + TN) / (TP + TN + FP + FN)

According to the parameters Precision, F1 score, and Accuracy, Table 4 compares the performance of several classifications including Nave Bayes, Support Vector Machine, Rule-based Sentiment Analysis, Decision Tree, Decision Support System, and Stacked DenseNet121 + attention-based MLP (Proposed). Comparing the total performance, the proposed method outperforms the other techniques currently in use in terms of accuracy.

Figure 5 provides a graphical representation of how different methods, including Naive Bayes, Support Vector Machines, Rule-based Sentiment Analysis, Decision Trees, Decision Support Systems, and Stacked DenseNet121 + attention-based MLP (Proposed), performed using various metrics, including Precision, F1 score, and Accuracy. It was established that the

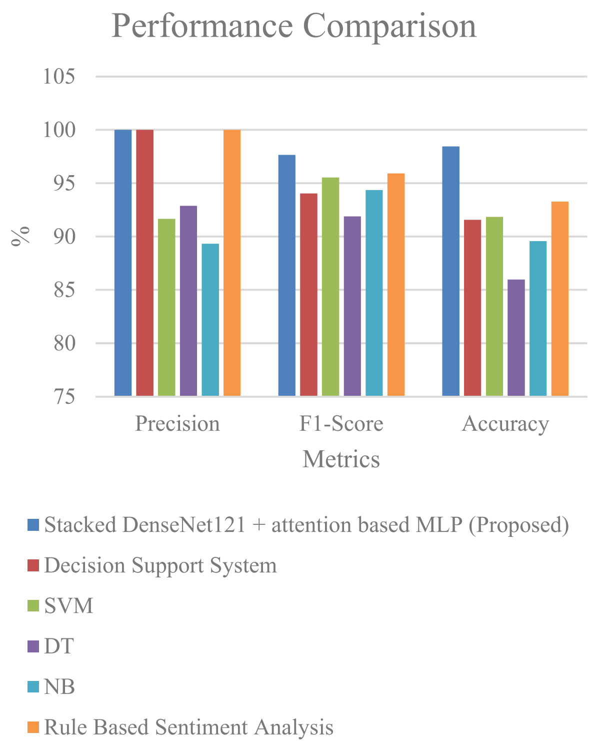

Figure 5. Performance results.

Table 5. Results of machine learning systems.

Classification techniques	Positive score	Negative score	Total	Positive score %
Support Vector Machine [5]	1687	438	2125	79
Logistic Regression [6]	1687	438	2125	79
Multinomial Naïve Bayes [8]	1687	438	2125	79
Random Forest [7]	1687	438	2125	79
Decision Tree [7]	1687	438	2125	79
Rule based Sentiment Analysis [17]	1761	294	2055	86
Stacked DenseNet121 + attention-based MLP (proposed)	1875	158	2033	92.22

suggested effort gave the best and most exceptional results.

Table 5 shows the results of the proposed approach Stacked DenseNet121 + attention-based MLP as well as the outputs of currently existing machine learning techniques as SVM, LR, Multinomial Naive Bayes, RF, DT, Rule-based sentiment Analysis. The proposed study produces the best results when compared to the existing methodologies. Out of 2033 reviews, the suggested approach has a positive score percentage of 92.22%.

Conclusion

In this approach, a deep learning-based framework was presented for relevant product recommendations in the social network. The main goal was to make it possible for computers to learn automatically without any assistance from humans, in order to control the operations. As a result, the social input dataset was originally pre-processed in this suggested work to remove noise. Fisher discriminant algorithm based on information was used for feature extraction. Then, a hierarchical agglomerative and attribute-based clustering technique was used to choose the features. Then, using stacked DenseNet121 classifiers, a prediction was created, and a product recommendation was made using attention-based MLP. In order to verify the effectiveness of the suggested system, the expected output was evaluated and the performance estimation was compared with conventional methodologies. When compared to the current approaches, the proposed study yields the best findings. The suggested strategy has a positive score percentage of 92.22% out of 2033 reviews. According to the analysis, the suggested system is more effective at providing better results for the recommendation of the pertinent goods. In future work, automatic extractions of explicit review or feedback from websites are to be added to improve the dataset quality for improved accuracy.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

[1] Wase K, Ramteke P, Bandewar R, et al. Sentiment analysis of product review. Int J Innov Eng Sci. 2018;3(5):8–13.
[2] Roul RK, Sahoo JK. Sentiment analysis and extractive summarization-based recommendation system. In: Computational intelligence in data mining. Singapore: Springer; 2020. p. 473–487.
[3] Praba MB, SRMIST C, Nadu T, et al. Educational website based on dynamic recommendation and filtering techniques. 2018;6:84–87.
[4] Dwivedi R, Anand A, Johri P, et al. Product based recommendation system on amazon data.
[5] Yesodha K, Anitha R, Mala T, et al. Product recommendation system using support vector machine. In: Advanced computational and communication paradigms. Singapore: Springer; 2018. p. 438–446.
[6] Bagde K, Patil J. Cold start product recommendation using social data and micro-blogging information.
[7] Malik S, Pandey Y, Kaur A. Recommender system: principles, techniques and application. Bloomsbury India; 2019. p. 98.
[8] Shaikh S, Rathi S, Janrao P. Recommendation system in E-commerce websites: A graph based approached. In: 2017 IEEE 7th international advance computing conference (IACC). IEEE; 2017, January. p. 931–934.
[9] Singla Z, Randhawa S, Jain S. Sentiment analysis of customer product reviews using machine learning. In 2017 *international conference on intelligent computing and control (I2C2)*. IEEE; 2017, June. p. 1–5.
[10] Senthilkumar MVR, Kiron CR, Vishnuvarthan M. A new approach to product recommendation systems. 2019.
[11] Deng X, Zhong Y, Lü L, et al. A general and effective diffusion-based recommendation scheme on coupled social networks. Inf Sci (Ny). 2017;417:420–434.
[12] Kiran MVK, Vinodhini RE, Archanaa R, et al. User specific product recommendation and rating system by performing sentiment analysis on product reviews. In:

2017 4th international conference on advanced computing and communication systems (ICACCS). IEEE; 2017, January. p. 1–5.
[13] Dadhich A, Thankachan B. Social & juristic challenges of AI for opinion mining approaches on Amazon & flipkart product reviews using machine learning algorithms. SN Comput Sci. 2021;2(3):1–21.
[14] Pothuraju LN, Pattanaik KK, Sahu R. A joint embedding technique for sequential recommendation. In 2020 *13th international conference on developments in eSystems engineering (DeSE)*. IEEE; 2020, December. p. 329–334
[15] Mongia S. Insights procured from the recommendation system based On user reviews. Turkish J Comput Math Edu. 2021;12(13):5587–5592.
[16] DWIVEDI R. Product based recommendation system on amazon dataset; 2021.
[17] Pote C, Makode P, Yadav A, et al. Sentiment analysis based product recommendation system. Int J Recent Adv Multidiscipl Top. 2021;2(6):155–159.
[18] Paranjape V, Nihalani N, Mishra N. A machine learning approach for item recommendation using SVD technique. In: 2021 IEEE 4th international conference on computing, power and communication technologies (GUCON). IEEE; 2021, September. p. 1–7
[19] Camacho LAG, Alves-Souza SN. Social network data to alleviate cold-start in recommender system: A systematic review. Inf Process Manag. 2018;54(4): 529–544.
[20] Portugal I, Alencar P, Cowan D. The use of machine learning algorithms in recommender systems: A systematic review. Expert Syst Appl. 2018;97: 205–227.
[21] Jing-Yu C, Ya-Jun W. Semi-Supervised fake reviews detection based on AspamGAN. J Artif Intell. 2022;4(1): 17–36.
[22] Kottursamy K. Multi-scale CNN approach for accurate detection of underwater static fish image. J Artif Intell. 2021;3(03):230–242.