Distributed Focused Web Crawling for Context Aware Recommender System using Machine Learning and Text Mining Algorithms

Venugopal Boppana[1]*, Dr. P. Sandhya[2]
Research Scholar[1], Associate Professor[2]
School of Computer Science and Engineering, Vellore Institute of Technology,
Chennai Campus, Chennai, India[1, 2]

***Abstract*—In today's world, Recommender System (RS) is the most effective means used to manage the huge amount of multimedia content available on the internet. RS learns the user preferences and relationships among the users and items. It helps the users to discover new interesting items and make use of different media types such as text, audio, video and images. RS can act as an information filtering model which can overcome the issues related to over-fitting and excess information. In this work, a new distributed framework named DAE-SR (Deep AutoEncoder based Softmax Regression) is introduced for context-aware recommender systems which focus on user-item based interaction and offers personalized recommendations. The proposed model is implemented in PYTHON platform. The dataset used for experimentation is Foursquare dataset. The performance of the proposed context-aware RS is beneficial to both the users and service providers. Its helps in decision making process and can offer relevant recommendations to users. The performance is evaluated in terms of various metrics such as accuracy, recall, precision and so on. From the implementation outcomes, the proposed strategy achieved good accuracy (98.33%), precision (98%), run time (1.43 ms) and recall (98.1%). Thus, it is proved that the proposed DAE-SR classifier performs better compared to other models and offer dependable and relevant recommendations to users.**

Keywords*—Recommender Systems (RSs); context-aware; softmax regression; deep autoencoder; multimedia information

I. INTRODUCTION

With the exponential development of products and services over the internet, individuals face issues related to data deluge as each individual have to make choices in a reasonable manner to save their valuable sources such as money and time [1-3]. Individuals possess admirable communication skills for expressing the ideas among each other in a proper form. The use of contextual information made the designers to increase the abundance of communication between human and computer interaction which enables the growth of intellectual applications. Now, RSs can be useful in various modern applications which expose the user towards a massive collection of items. These structures usually offer a list of recommended items to the user in which each individual can predict or prefer an item [4-5]. These procedures are helpful to the users to choose the most suitable items and may ease the process of finding the desired items in the recommendation list.

The web crawler has turned out to be one of the key innovations for clients to naturally get data from assigned locales. The customary web crawler innovation has uncovered a few issues, for example, low material exactness because of basic sifting conditions regarding crawling themes, low productivity because of substance duplication and long website page refresh time [6-7]. Web crawler innovation is acquainted as a method to productively obtain and process data, and it includes different research territories, for example, distributed computing, data recovery, machine learning and web application [8-9]. Web crawlers are for the most part used to make a duplicate of all the visited pages for last handling. Additionally, crawlers can be utilized to assemble particular sorts of data from website pages, for example, gathering email addresses [10-11]. Quick fetching of relevant pages for the given topic with very less searches is possible with focussed crawler.

RS aims to seek the satisfaction of the individual/user via personalization and chooses the most appropriate services, items and provides it to the users considering the details such as user profile, opinions, preferences, purchase history, relationships with clients and interactions with services and products offered. The dual main entities of the RSs are the users, items which learn to create predictions by learning the parameters which minimize the loss among the actual and predicted preferences. Moreover, the RSs can be classified into 3 types such as CF (Collaborative Filtering), CB (Content-based) and HR (Hybrid Recommendation). Among these three categories, CB model gain advantages over the traditional RS system by exploiting the features of context-information as user activity, time, and location.

The integration of RS with context information influences the value of these models by enhancing the possible relevant recommendations in accordance to the varying user requirements. The CEP (Complex Event Processing) is integrated with RSs to provide personalized recommendations and allows the business users to put on event-driven rules for controlling and filtering the input parameters of recommendation engines [12-14]. The concept of context-based RSs minimizes the gap among the information models and users, therefore these information models can actively realize the user's context and offer better experience to users [15-16]. Recommender systems (RSs) are a subclass of data filtering frameworks that look to anticipate the rating that a client would provide for an item.

Recommender frameworks have been connected to variety of uses, e.g., films, music, news, books, look into articles, seek inquiries, social labels, and money related administrations. As a rule there are three different ways to plan recommender frameworks, i.e., cooperative separating, content-based filtering [17], and the hybrid filtering [18]. Recommender models have attracted consideration in both industry and academia. Such frameworks help to oversee data over-burden via assembling the data self-sufficiently and fitting it to singular interests [19]. A recommender framework would have the capacity to catch the clients' preferences. Social Recommender Systems (SRSs) or Recommenders for the Social Web demonstrate the client's preferences by utilizing the data he or she and their companions have distributed in online informal communities [20-21].

A. Motivation

Recommender Systems gained benefits by solving the problem of handling abundant data in the internet. Individuals plan to make effective and smart decisions on the huge number of available choices in various fields such as: a place the person wishes to visit, a movie the person likes to watch, a book he/she wants to read etc. The main motive behind this generation of recommender system is to offer the intelligence that benefits both the consumers and organizations. However, there is an emerging trend of using contextual information to afford meaningful personalized recommendations. The context enhances the basic interaction of user-item to realize recommendations of higher quality. Moreover, recommendation using content-based framework effectively use the information content describing the item or user such as item description or user profile to execute accurate recommendations. Also, content-based (CB) recommenders are more suitable with sparse datasets and may not suffer from the cold-start issue. As the textual data is used to represent the item description and user profile, this content based context-aware model make use of the common evaluation metrics to measure the similarity between the contents. The recommendations based on contextual information emphasize that the choice made by individuals is liable on certain context, rather than being invariant of it. Sometimes, the same users might select diverse services/products under a dissimilar context. Hence, this condition motivated the authors to develop a precise prediction of preferences by consumer depending on the appropriate contextual data which is being combined with the recommendation method.

B. Contributions

In this work, context-aware approach is developed by utilizing the contextual user information, in order to produce more accurate recommendation according to user's preferences. This recommendation model makes use of the information from the dataset in order to create user recommendations on basis of given number of choices. The main novelty of the proposed DAE-SR recommendation model is to develop a context-aware/content-based RS that estimates the characteristics of restaurant and relevant context factors against the user preference model to create top recommendations for restaurant names based on the similarity score calculated. These scores can be used later to sort out the top restaurants for each user according to their expectation. Our proposed (DAE-SR) RS adopts a user preference model by using features of liked/visited restaurants by the user and takes relevant context factors, e.g., restaurant ratings, location, current weather, and time of the day. Most of the existing research works use CF based recommendation because of its higher performance accuracy. Since, CF approach is more effective and reliable one of the major drawbacks reported was the higher scalability. Hence, a new recommendation framework is presented in this work based on context-aware which improves the quality of recommendations. The major contributions of this paper are summarized as follows:

- To introduce a distributed framework for context-aware RS based on text mining and machine learning algorithms in order to afford quick and correct suggestions to user on good restaurants name.
- To present the content-based context-aware RS to create an item list that reveals the user's preference and minimizes the error in the classification process.
- To develop a hybrid DAE-SR based recommendation on restaurant names which is trained with the meta-heuristic algorithm called JO (Jaya Optimization) algorithm which has good ability of exploration and exploitation.
- The three similarity measures such as Dice's, Jaccard and Cosine similarity coefficients are used to compute the similarity of the content and to obtain high accuracy in receiving top recommendations.

C. Organization

The paper is organized as follows. In Section I, a brief introduction about the topic is given. Section II gives the works related to our proposed algorithm and Section III gives the proposed methodology. The result of the proposed algorithm is provided in Section IV and the entire work is concluded in Section V with future suggestion and references.

II. Related Works

The main motivation of related works section is to explain about the importance behind each research. In this this section, the discussion about the features of existing approaches, its applications, limitations, metrics and datasets used is clearly described in Table form.

Afolabi *et al.* [22] have presented a semantic web content mining scheme for RS in shopping by online. Through a web crawler, the web textual dataset was gathered built in Java. The combination of an existing ontology and a developed ontology was utilized for the textual data's semantic preprocessing. Next, the recommendation was created by the Naïve Bayes algorithm. The large dimensions of attributes were handled by the annotations to ensure the preprocessing models optimized semantic understanding.

Amato *et al*. [23] presented a technique SOS: A multimedia RS for Online Social systems. In this viewpoint, the RS was introduced for supporting the browsing of information accumulations and helping users to discover what they truly required. A new RS for big data applications was introduced in this research, which was able to deliver recommendations based on the interactions between users.

A hybrid learning based recommender framework was proposed by Tarus *et al.* [24]. In this paper, hybrid information based RS was introduced on the basis of sequential pattern mining (SPM) and ontology. The domain knowledge about the learning and learner resources was represented and modelled by the ontology, although the SPM algorithm determines the learners' sequential learning patterns.

RS with Linked Open Data (RS-LOD) and matrix factorization process (MF-LOD) was proposed by Natarajan *et al.* [25] for reducing the problems of cold- start and sparsity. The cold start issue was solved by the RS-LOD approach, and the data sparsity problem was eliminated by the MF-LOD approach. The hidden feedback data was elongated the user vector, and the similar semantic items were elongated each item. The collaborative filtering RS's precision was enhanced by the Semantic features from LOD.

Iqbal *et al.* [26] proposed a context-aware kernel-based recommendation scheme that creates the design over the user-item matrix of context rating. The importance of contextual information was considered for this recommendation. The KCR (Kernel Context Recommendation) algorithm was accurate, flexible and scalable enough to include the various contexts and was utilized for creating the practical recommendations.

A Preference Relation based PMF framework for RSs was proposed by Pujahari and Sisodia [27]. User's PR was taken as input, and the recommendation was also generated. The user's and item's neighborhood data were integrated into the design. The users' preferences towards items were obtained by the Probabilistic Matrix Factorization (PMF) system. This approach was introduced for Collaborative Filtering (CF) in RS. Aghdam [28] proposed a hierarchical HMM (hidden Markov model) in which the changes were identified in the preferences of user on time through modelling the user's latent context. The user as a hidden Markov process was modelled by the user-selected items.

Sangaiah, A.K *et al.* [30] developed a new technique with the help of ML (Machine Learning) methods to conserve the confidentiality of PBSs (Position Based Services) roaming users. The emergence of the three-phase procedure was used for the identification of PBS user position by means of combining k-NN (k-Nearest Neighbours) and DT (Decision Trees). The destination of the user was estimated by means of the HMM (Hidden Markov Models) and the position track sequence. The service policy followed was the mobile edge computing which was useful for ensuring the PBS timely delivery.

Sangaiah, A.K *et al.* [31] presented an energy-aware model (EAM) named green adversary for the application of smart industrial environment. The proposed model conserves the information and position confidentiality whereas the existing models make use of the cyber-physical systems by mutually using the software and hardware parts in which the energy consumption was minimized. However, this limitation was tested on the recommended approach related to the CPS (Cyber Physical Security) smart industrial applications and has removed the severe effects and also tested on CPS real-world applications.

Sangaiah, A.K *et al.* [32] presented whale optimization algorithm (WOA) to obtain optimal solution for network issues. WOA was powerful heuristic algorithms that boosted the exploration process and prevented bottleneck. However, the main drawback of this algorithm was a low convergence rate. Likewise, Bat algorithm (BA) was a novel meta-heuristic algorithm presented by Sangaiah, A.K *et al.* [33] to achieve optimal solutions in terms of sustainability and quality. It has the ability to monitor and control the network activities. Though, the performance of BA for huge data was not efficient. It was considered as a major disadvantage in BA.

Gupta Garima and Rahul Katarya [36] presented an Ensemble Particle Swarm Optimization (EnPSO) approach which was considered as an AutoML (AutomatedML) to minimize the model selection complexity. The presented EnPSO optimizes the recommendations intelligently using the identification of best ensemble model. The dataset used for testing was MovieLens in which the recommendations were improved with the AutoML system. The accuracy of the model was verified with the RMSE metric with lower error as 0.918 compared to the baseline IBCF (Item-Based Collaborative Filtering) with Singular Value Decomposition (SVD) having 0.961 higher errors. The main challenge was to find an effective way to reach the performing scheme with the vast search space. The limitation of the suggested approach was that EnPSO cannot provide effective performance gains within the constrained time for some cases.

Katarya Rahul, and Yamini Arora [37] developed a novel scheme called CapsMF (Capsule Networks Matrix Factorization) for product RS. The DNN (Deep Neural Network) text analysis architecture was enhanced with the Bi-RNN (Bi-directional Recurrent Neural Network) for the representation of text descriptions in robust form. The integration of Probabilistic MF with the DNN has the ability to generate recommendations in improved manner. The accuracy of the system was analysed with MAE (0.8878) and RMSE (1.157) shows improved outcomes. The major limitation observed was the training time was quite higher with the Caps Network during the experimentation process.

Katarya Rahul [38] presented a reliable RS model using the improved CF technique. A new self-devised method named P-distance algorithm was used to predict the ratings and acts as a dual-level filtering procedure to identify the nearest neighbours. The experiments were tested on 3 datasets such as Douban, Movielens and Jester. The performance was evaluated with MAE and RMSE metrics to show the effectiveness of the proposed algorithm. Sparsity was the major problem identified with this reliable RS.

Katarya Rahul [39] developed a hybrid RS utilizing KMC (K-means clustering) with ABC (Artificial Bee Colony) optimization procedure. The hybrid KM-ABC approach offers accurate movie predictions by considering the user ratings.

Various metrics such as accuracy, recall, precision, MAE were evaluated for verifying the system performance on the Movielens dataset. The experimental results on the dataset indicate that the proposed KM-ABC based recommender system has gained improved performance regarding reliability, accuracy, and personalization for the recommendation of movies. The drawback associated with this system was that only user ratings were considered and the other user characteristics were not considered.

Yadav *et al.* [40] presented a review on RS using FL (fuzzy logic). With the combination, with fuzzy logic three categories were reviewed such as content-based, memory-based and model-based systems. FL was utilized to assign the fuzzy values to the user items and represents the preference of user in robust form. The Fuzzy recommender was used to compute similarity using the fuzzy values rather than the user-based crisp data. This review offers significant knowledge about the various fuzzy based RS models. The fuzzy logic imitates the decision making process of humans which cover all the intermediate values.

Katarya Rahul [41] developed an improved RS utilizing dual-level MF (matrix factorization). This improved framework incorporates product ontologies for learning the items recommendations. The proposed matrix – factorization based RS initiates each item by generating similarity measures with each other. The similarity was created with the help of genres and a matrix was created using the given user ratings, so that the user gets his corresponding recommendation. Sparsity was the major problem identified with this reliable RS. The dataset used for experimentation was MovieLens dataset and the metrics used for evaluation were MAE and RMSE.

Katarya Rahul and Om Prakash Verma [42] introduced a web-based RS which was mainly based on successive information of user's navigation on web pages. FCM (Fuzzy C Means) approach was integrated to receive top-N clusters. Weights were calculated for each of the page category and the top page recommendations were predicted for the targeted user. The experimentation was performed on the real-world MSNBC dataset. Compared to other clustering approaches the proposed FCM model has gained 33%accuracy. The proposed RS required lower computation speed of 402 seconds whereas the other methods show similar as well as enhanced outcomes.

Katarya Rahul and Om Prakash Verma [43] presented a hybrid model for music recommendations. The proposed HMRS (Hybrid Music RS) represents the combination of both collaborative-content features. The algorithms such as BF (Bellman-Ford) and DFS (Depth First Search) were utilized with the multi-layered context graph. Here the similarity score was computed using PC (Pearson Correlation) coefficient. The overall results were optimized with the PSO algorithmic procedure. The music dataset used was Last.fm and the accuracy of the system was evaluated with recall metric on Top-N recommendations.

Katarya Rahul *et al.* [44] developed a location-based RS framework with the help of improved RW (Random Walk) model. Both the structural as well the attribute properties were utilized to recommend the location in SNs (Social Networks). The computations of accurate recommendations on locations were done with the improved RW model. The experimentation was conducted with three real-world datasets such as Gowalla, Brightkite and Foursquare. The accurateness of the system was evaluated with precision, F1 score and recall metrics.

Katarya Rahul *et al.* [45] presented a HF (Hybrid Filtering) approach in analysing the user behaviour in content-aware RS. The dual filtering strategies used are the pre-filtering and post-filtering models. The HF model eliminates the recommended items which has the lesser chance of relevance. Based on the significance of the contextual attribute both the filters were combined which forms a hybrid model. This HF approach minimizes the sparsity problem also it was quicker than the post-filtering method. The experimentation was performed on 50 movies that were selected equally from Movielens dataset and the precision calculation determines the accurateness of the system. Sanchez et al. [46] presented a RS for food delivery on the basis of number of orders. Nearest-neighbour (NN) algorithm was used for the evaluation of users preferred restaurants and buying patterns. Teixeira et al. [47] introduced a diabetic friendly restaurant recommendation using MAS (Multi-Agent System) for multi-criteria decision making. The users with diabetic illness can be benefitted and can find out suitable nearby restaurants. Table I represents the review on various recommendation systems.

A. *Research Gaps*

Most of the works developed in the context aware recommender system concentrates only on improving the accuracy of recommendation. But, the major purpose is to improve the recommendation accuracy when considering the design objectives such as a user's item's context. Moreover, the main challenge of a RS is to generate a significant recommendations using context based rating of user and item information. Hence, the recommender system should be accurate, flexible and that should predict the contexts accurately. Moreover, some of the recommender systems are providing the improved accuracy but not providing the rating of contexts. Some of the recommendation systems shows decreased computational complexity but fails in the accurate prediction and also results in the difficulty of generating recommendations.

The running time of the previous RS is noted to be higher while providing the recommendations which depend on the previous preferences of users. The preference may sometimes change based on time because of the mood change of users otherwise the change of contexts. Numerous earliest techniques are presented based on the statement of every user having the static pattern. But without these changes, the recommendation does not equal the user's choices. The context aware dependent recommendation systems are dealing with this problem through the contextual data. The contextual data usage is difficult because of attaining all contextual data for processing. Moreover, the different types of contexts used in the RS extend its dimensionality. Hence, to overcome all these issues, the authors of this work presented a Distributed Focused Web crawling with Machine Learning and Text Mining algorithms for the effective context aware recommendation system.

TABLE I. REVIEW ON VARIOUS RECOMMENDATION SYSTEMS

Author and Reference	Technique Used	Dataset Used	Applications	Metrics	Limitation
Afolabi *et al.* [22]	Semantic Web Mining + NB classifier	Wed textual data collected using Web crawler	Product Recommendation	Precision, Recall, F-score	Results are not optimal.
Amato *et al.* [23]	SOS	YFCC100M	RS for big data applications	MAE, RMSE	Information overload
Tarus *et al.* [24]	Hybrid recommendation framework	Real world dataset	e-learning	MAE, Precision, Recall	Machine learning techniques are not used.
Natarajan *et.al.* [25]	RS-LOD	Netflix and MovieLens datasets	Social network, e-commerce	MAE, RMSE, Precision, Recall, F1-score	Machine learning and Deep learning techniques are not used.
Iqbal *et al.* [26]	KCR	LDOS-CoMoDa, DePaulMovie datasets	Movie Recommendations	RMSE, F1-Measure	Do not consider the context under which ratings are provided.
Pujahari and Sisodia [27]	PMF	MovieLens 1M and MovieLens 20M datasets	Movie Recommendations	MAP, NDCG	Along with user preferences there will be presence of side information
Aghdam [28]	Hierarchical HMM	Last.fm and Netflix datasets	Music and Movie Recommendations	Precision, F-measure, Recall	Sparsity problem
Sangaiah, A.K *et al.* [30]	PBS	Real-time dataset	Industrial informatics	Search space	Confidentiality issues
Sangaiah, A.K *et al.* [31]	EAM	-	Real-world applications	Query Time	Confidentiality issues
Sangaiah, A.K *et al.* [32]	WOA	Test data	Real-world applications	Convergence	Computational complexity
Sangaiah, A.K *et al.* [33]	BA	-	Real-world applications	Distance, Priority	Computational complexity
Gupta Garima and Rahul Katarya [36]	EnPSO	MovieLens dataset	Movie Recommendation	RMSE, MSE, MAE	Vast search space
Katarya Rahul, and Yamini Arora [37]	CapsMF	AA (Apps for Android), AIV (Amazon Instant Video)	Product Recommendation	RMSE, MAE	Higher training time
Katarya Rahul [38]	P-distance algorithm	Douban, Movielens and Jester	Movie Recommendation	MAE, RMSE	Sparsity
Katarya Rahul [39]	KM-ABC	Movielens	Movie Recommendation	accuracy, recall, precision, MAE	Only user rating is considered
Yadav *et al.* [40]	FL	-	Review on various FL	-	-
Katarya Rahul [41]	Two-Level MF	Movielens	Movie Recommendation	RMSE, MAE	Sparsity
Katarya Rahul and Om Prakash Verma [42]	FCM	real-world MSNBC dataset	Acquiring user's social information	Accuracy, Time	Reliability issues
Katarya Rahul and Om Prakash Verma [43]	HMRS	Last.fm	Music Recommendations	Recall	Run time complexity
Katarya Rahul *et al.* [44]	Improved RW	Gowalla, Brightkite and Foursquare	Recommend location in social networks	Precision, recall, F1-score	Computational Intelligence and Data mining techniques are not used.
Katarya Rahul *et al.* [45]	HF	Movielens	Movie Recommendation	Precision	Time complexity
Sanchez et al. [46]	NN	DeliveryFood application	Food-delivery	No. of neighbours (k) and percentage of orders (pmin)	Run time complexity
Teixeira et al. [47]	MAS	Real-time dataset	Diabetic-friendly restaurant recommendation	-	Reliability issues

B. Research Questions

The recommended system aims to execute certain tasks in assisting smart recommendations on restaurants. In this work, context aware-RS have played a most significant role in generating recommendations. Thus, some of the following research questions presented are:

- How are the contexts collected?
- Which contexts are used in smart learning?
- What data mining and recommendation techniques used to processes the contexts?
- What are the recommended activities?

The existing approaches utilized for the recommendation system was discussed in Section II with its research gaps. To avoid the issues presented in the existing RS, a new approach is proposed and the research questions are framed which is solved in this presented work (Section III).

III. Proposed Methodology

The recommendation system uses one of the exceptional pieces of the researches, which can be used to offer better solution for challenges also it uses supplementary refined machine learning algorithms for example, ANN. To analyse the performance of the Context-aware recommendation systems a combination of different approaches is used. The initial attempt is to model the criteria rating, approximately powerful machine learning approaches are used for this process. In order to enhance the prediction accuracy the deep learning model is used in this work. This paper make use of deep auto encoder and Softmax regression trained (DAE-SR) with a Jaya optimization (JO) algorithm for providing effective recommendations. We selected the Jaya optimization algorithm, since it has a high tendency to high convergence rates, low training time, and escape local minima compared to other optimization algorithms.

Fig. 1 shows the overall workflow diagram for the proposed strategy. For the web data acquisition, crawler is a significant tool. However the frequent updates of the data sources, distribution channels and web data structures are resulted in high costs of crawler program maintenance and development. From the raw web data the textual information's keywords are extracted by means of measuring keyword comprehensive weights via informative features. Textual information's keywords reflects the ideal way for data mining also establishes the knowledge representation model, after that constructs the index library. At last, for the deep machine learning strategies build a prediction model to obtain relevant recommendations. The prediction is performed based on the strategy named as DAE-SR to obtain effective results. The textual information pre-processing consists of following phases. They are:

- Information Collection and Processing.
- Pre-processing.
- Similarity Calculation.
- Recommendation Prediction.

A. *Information Collection and Processing*

For collecting the news the web crawler provide important theoretical support. Web crawler is a central part of search engines. The use of web crawler is not difficult when getting the news from the internet. Implementing a new crawler and not extending the existing crawler module by the user is the major problem with crawler. It is used to rescue the web pages also attach them into the local source. In a centralized location, collecting and processing the entire contents of web is the general-purpose of web crawlers. Every day a huge volume of web pages are frequently added and data is persistently changing due to the rapid emergence of internet. On the basis of event pattern or event procedure rule, the CEP is said to be a real-time data processing methodology. It is used to retrieve the high level knowledge from the large extent of data. It can also be utilized to analyze the trends, track the data from multiple sources, and patterns. The events that happened in past, are used in any order which can be allowed by CEP.

B. *Textual Information Pre-processing*

Pre-processing strategies named as stop words removal, word extraction, stemming and also term frequency–inverse document frequency (TF/IDF) are the important approaches in textual information pre-processing. Based on the words existing in the user opinion feedback reviews, the words are tokenized, count of frequency of word and stem are done by using the pre-processing strategies. Fig. 2 shows the process of pre-processing.

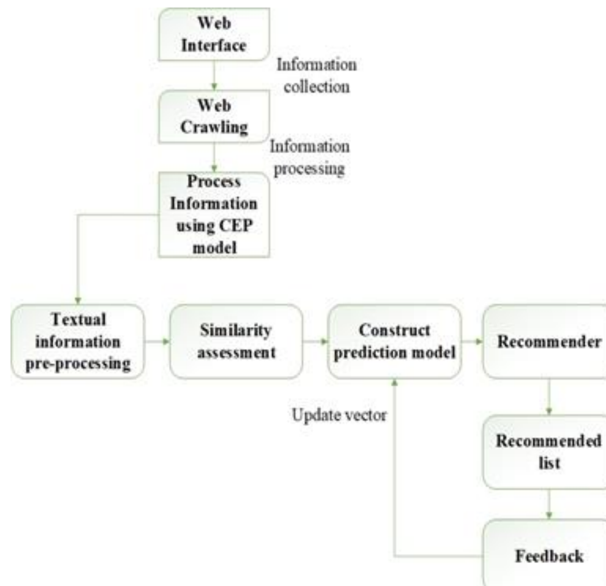

Fig. 1. Workflow diagram for proposed methodology.

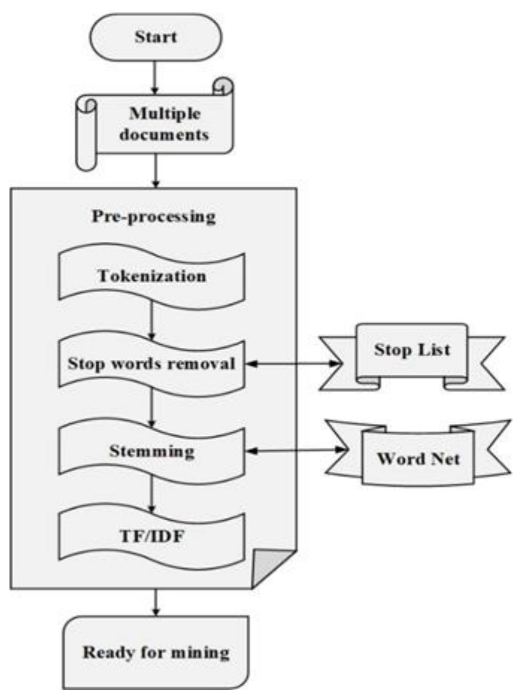

Fig. 2. Model of text-mining pre-processing.

Pre-processing plays a vital role in text mining (TM) applications. Pre-processing is the first step, which involves the combination of other steps such as token extraction, stemming, removal of stop words, and TF/IDF tasks.

1) Word (tokenization) extraction: The procedure of converting the sensitive data into tokens is called as tokenization. Here the significant amount of text structure is reduced to a word, phrase, or symbol. Using NLP (Natural Language Processing), the word in the document set *D* is tokenized.

2) Stop words: Pronouns, prepositions and articles are generally present in documents. These words may not add much meaning and considered as useless. Some of the stop words used are ‘with, an, is, the, has, in, be and, at’ etc. In text mining applications stop words are not considered as keyword, so it is eliminated from the document.

3) Stemming: The process of extracting the base form of words by eliminating the affixes and stem the root of a word into the original word is termed as stemming, example; programs, programming, programmed $\rightarrow$ ‘program’ etc.

4) Term frequency–inverse document frequency computation (TF-IDF): For text mining, TF-IDF is commonly applied as a weighting factor. A word how important to a document is given in this method; it is also known as numerical statistics.

a) Term weighting systems: Term weights schemes i.e. Document Frequency (DF), Term Frequency (TF) and Inverse Document Frequency (IDF) are frequently used. For the arithmetic representation of collection of web pages, the VSM (Vector Space Model) is the best widely held and proficient approach. Each web page is deliberated as a vector of terms in VSM model such that $d = (t_1, t_2, ..., t_n)$, and an equivalent weights vector $w = (w_1, w_2, ..., w_n)$, where $w_{1,} w_2, ...w_n$ the weights of $t_1, t_2, ..., t_n$ are correspondingly, based on the used term weighting system. Table II gives the example of two dimensional based VSM model.

b) Term frequency (TF): In the web page, TF is related through the weight of a definite term. Normalized TF is measured in this paper. By finding the amount of raw frequency of the term t in web page w is calculate weight of term t is displayed in Eq. (2). The square root of the addition of the square of frequencies of all terms in the web page is known as Euclidean norm.

$$TF_{t,w} = \frac{fr_{t,w}}{\sqrt{\sum_{t=1}^{n} fr_{t,w}^{\ 2}}} \quad (1)$$

Table III shows the example of collection of web pages, in this highest TF value is mentioned as bold and underlined cells is the second highest value in the matrix. As shown in Table III, on page 3 the term ‘payment’ has the highest *TF* value (0.7071). The term ‘payment’ is placed two times in page 3 and 4. The TF value is normalized to webpage’s Euclidean norm.

TABLE II. TWO DIMENSIONAL MATRIX BASED VSM MODEL

Web Page \Term	**account**	**service**	**banking**	**payment**	**cheque**	**insurance**	Page Length*
Page 1		1		1		1	3
Page 2	1		1		1	1	4
Page 3	1	1		2	1	1	6
Page 4	1		1	2	2	1	7
Term Frequency **	3	2	2	5	4	4	
Page Frequency ***	3	2	2	3	3	4	

* The number of distinctive terms in the web page
** Summation of term occurrences in entire web pages.
*** The number of web pages in which the term appears.

TABLE III. COLLECTION OF WEB PAGE AND HIGHEST TF VALUE

Web Page \Term	**account**	**service**	**banking**	**payment**	**cheque**	**insurance**
Page 1		0.5773		0.5773		0.5773
Page 2	0.5000		0.5000		0.5000	0.5000
Page 3	0.3535	0.3535		0.7071	0.3535	0.3535
Page 4	0.3015		0.3015	<u>0.6030</u>	<u>0.6030</u>	0.3015

c) Document frequency: The number of web pages within the collection is represented as DF (weight of term *t*), here the term *t* is found. It is a global weighting term.

$$DF_t = \sum_{p=1}^{N} \begin{cases} 1 & t \in p \\ 0 & t \notin p \end{cases} \quad (2)$$

TF–IDF: The terms within the collection of web pages are denoted by TF-IDF, which is an effective ranking measure. It also reproduces the statement that less frequent word in the collection is more significant word in web page and viz. The dot product of TF and IDF of a term is denoted as TF-IDF. Table IV gives the example on the basis of TF–IDF weight scheme.

$$TF - IDF_{t,w} = TF_{t,w} \times IDF_t \quad (3)$$

Where $IDF_t = \log\left(N / DF_t \right) + 1$

The term 'payment' appears three times in the document, so it has highest TF-IDF value shown in Table IV. Also, the term 'banking' and 'service' appears twice in the collection.

C. Similarity Calculation

After the text pre-processing the opinions are grouped together according to their similarities, while the huge number of reviews are received. Similarity measurement is necessary in text mining approaches. Similarities among the string are calculated by similarity measures. Dice's coefficient, Jaccard similarity coefficient and cosine similarity are defined below:

1) Dice's coefficient: This metric is utilized to measure the documents similarity, and the inventive formula is given in Eq. (4).

$$SQ = \frac{2|B \cap C|}{|B| + |C|} \quad (4)$$

Where, $|B|$ *and* $|C|$ signifies the number of terms present in documents. *SQ* represents the similarity quotient.

2) Cosine Similarity (CS): Two vectors of an inner product space similarity are calculated by this method, cosine of the angle is measured. Cosine of 0° value is 1, for some other angle it is <1. Mostly, in positive space cosine similarity is used, its results efficiently limited in (0, 1). Cosine Similarity [CS or *cos(θ)*] can be computed using the following eqn. 5 & 6,

$$CS = cos\,\theta = \left\{ \frac{Dot\ product(V,M)}{\|V\| * \|M\|} \right. \quad (5)$$

$$cos\,\theta = \frac{\sum_{i=1}^{n} V_i M_i}{\sqrt{\sum_{i=1}^{n} V_i^2}\sqrt{\sum_{i=1}^{n} M_i^2}} \quad (6)$$

Where, V_i and M_i signifies the components of vector V and M.

3) Jaccard Similarity (JS) Coefficient: JS is utilized to match the diversity and similarity of model sets. The other names are Jaccard coefficient/ index. The size of intersection divided by the size of union of the sample sets is known as Jaccard coefficient and computes the similarity among finite set of samples as shown in Eq. (7),

$$JS(B,C) = \frac{|B \cap C|}{|B \cup C|} \quad (7)$$

Where, B and C represent the dual sets.

D. Recommendation Prediction: Hybrid DAE-SR Model

The performance of proposed recommendation model is investigated via neural systems to analyse the textual item content. In this paper, we proposed Softmax regression-based deep auto encoder network (DAE-SR) model trained with Jaya Optimization (JO) algorithm. Compared to the existing models activated in RS, in which the DAE provides better outcomes because of its high capability to reconstruct the inputs. Meanwhile, the softmax regression is utilized for classifying the ratings. We picked the JO algorithm since it has a high propensity to escape nearby minima, low preparing occasions, high assembly rates contrasted with other enhancement calculations, which is utilized to produce improved recommendation prediction.

TABLE IV. EXAMPLE ON THE BASIS OF TF–IDF WEIGHT SCHEME

Web Page \ Term	**account**	**service**	**banking**	**payment**	**cheque**	**insurance**
Page 1		0.7510		0.6494		0.5773
Page 2	0.5624		0.6505		0.5624	0.5000
Page 3	0.3976	0.4599		**0.7954**	0.3682	0.3535
Page 4	0.3391		0.3922	0.6783	0.6783	0.3015

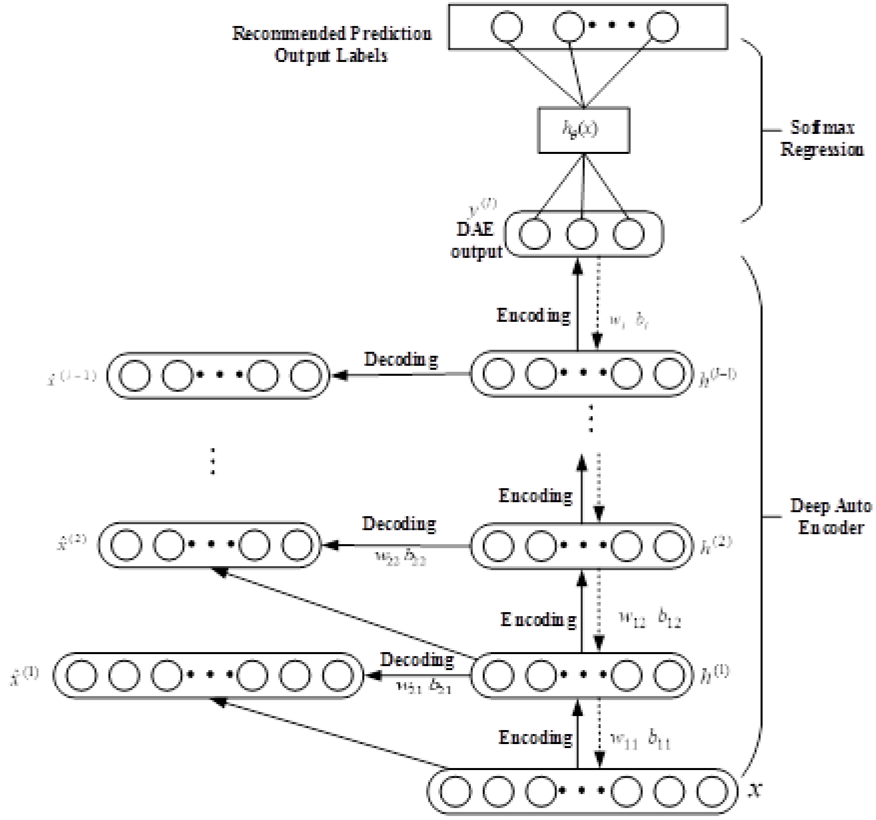

Fig. 3. Architecture of DAE-SR model.

Fig. 3 represents the architecture of DAE-SR model. It can be seen that DAE and softmax regression can be perfectly linked as a soft-hybrid model, which can make full use of labelled instances and achieve supreme accuracy with Jaya optimization algorithm. Here, we mainly explain the hybrid model of DAE-SR for recommended prediction. The deep auto encoder which is the popular deep learning model can learn discriminative features to produce autonomous distributed depictions of contents of items and users. Table V signifies the details of hyperparameters.

TABLE V. HYPERPARAMETER DETAILS

Hyperparameter	Value
No. of hidden layers	3
Batch size	16
Number of epochs	50
Activation	softmax
Learning rate	0.001
Optimizer	Jaya Optimization (JO)
Epoch	100

A DAE is said to be a deep neural network, which is fabricated via stacked Auto encoders. Pre-training that stacked AE from the bottom to top is a major phase of training a DAE. The result features of the hidden layer designated as input of the top-layer of the DAE, when the constraints of bottom layer of the DAE are decided. Through the Jaya optimization algorithm the procedure of the error back propagation is originated to fine-tune the constraints of the complete network. The number of layer fix as l also the function of objective can be displayed as in Eq. (8):

$$s\left(w_l, w_{i,k}\,|_{k=1}^{l-1}, b_{i,k}\,|_{k=1}^{l-1}\right) = \underset{w_l, w_{i,k}, b_{i,k}}{\arg\min} \frac{1}{2N} \sum_{i=1}^{N} |y_i - g_l(f_l(h_i^{l-1}))|^2 \quad (8)$$

Where, activation esteem of $(l-1)^{th}$ hidden layer indicated by $h_i^{l-1} = f_{l-1}(f_{l-2}(\cdots f_1(x_i)))$, and y_i signifies the label of x_i. W_l is the final layer weight, $w_{i,k}$ and $b_{i,k}$ represent the weight and bias of k^{th} layer, correspondingly. Moreover, the constraints of DAE are restructured as:

$$w_l := w_l + \Delta w_l = w_l - \mu d^l h^{l-1} \quad (9)$$

$$w_{1,k} := w_{1,k} + \Delta w_{1,k} = w_{1,k} - \mu d^k h^{k-1} \quad (10)$$

$$b_{1,k} := b_{1,k} + \Delta b_{1,k} = b_{1,k} - \mu \sum_{j=1}^{R} d^k \quad (11)$$

Where, the value $d^l = (h^l - Y)h^l(1-h), d^k = w_{k1}^t d^{k+1}(1-h^k)(when\ k < l)$, learning rate is represented as μ and size of input features denoted as R. The repetitions of $w_{1,k}$ and $b_{1,k}$ are performed till the function of objective reaches the max-epoch/iteration. Hence, the output of DAE is given to the SR model for further processing. Softmax Regression (SR) is a single-layer neural network used to evaluate the conditional probabilities linked with all the possible number of classes. A logistic regression display is the most straight forward type of a neural system. It comprises of an info layer with numerous traits and a bias unit, and just a single yield layer, or class. It is basically a double classifier. The Softmax Regression calculation applies binary strategic regression to different classes without a moment's delay. Let $D = \left\{ \left(z^{(1)}, y^{(1)}, \ldots\ldots, (z^{(n)}, y^{(n)} \right) \right\}$ be the training set, $z^{(i)}(i \in \{1,2,\ldots.,n)\}$ signifies the training data, $y^{(j)}(j \in \{1,2,\ldots\ldots,n\})$ signifies the label of training data. Given a test instance z, to evaluate the posterior probability $p(y = k \mid z)$. Softmax regression consists of input, classifier and output. A Softmax function is defined as the following Eq. (12):

$$P\left(y = \frac{j}{z^{(i)}} \right) = \phi_{soft\max}\left(z^{(i)} \right) = \frac{e^{z^{(i)}}}{\sum_{j=0}^{k} e^{z_k^{(i)}}} \tag{12}$$

Where,

$$z^{(i)} = w_0 x_0 + w_1 x_1 + \ldots. + w_m x_m = \sum_{l=0}^{m} w_l x_l = w^T x \tag{13}$$

The Softmax function computes the probability that the training sample x_i belongs to class j given the weight and net input z_i. The process of finding the optimal weight parameters is the need for training DAE which minimize the objective function, where the models are trained by JO algorithm [29]. Compared to other optimization algorithms, this JO algorithm is said to be a powerful global optimization approach which can be used to solve the constraint as well as unconstrained problems using the benchmark function. Some of the advantages of JO algorithm are: It does not have any specific algorithm parameters to conduct the real computational experiments and has better algorithm convergence. The supreme reason to choose JO is the victorious nature of this algorithm which makes it more powerful than other evolutionary and swarm intelligence algorithms. The main objective of JO algorithm is, for a particular problem if the solution is once achieved the optimal result is reached simultaneously thereby avoiding the worst result. JO algorithm is most pretentious in application viewpoint and endeavours extraordinary accomplishments to discover the genuine solution successfully. Compared to other optimization procedures this JO algorithm utilizes only dual parameters such as number of iterations and size of population. An imperative benefit of using this algorithm is that it minimizes the time required for the optimization process and ignores the endeavour of varying constraints. Table VI specifies the pseudocode of Jaya Optimization.

TABLE VI. PSEUDOCODE OF JAYA OPTIMIZATION

Jaya optimization algorithm procedure

Initialize: population size P, number of iterations I
Initialize P solutions randomly
Find $f(z_i)\ \forall i = 1,2,\ldots,P$
Sort the population as Z_1 -best solutions, Z_P -worst solutions
t=1
while $(t \le I)$ **do**
  for $i = 1,\ldots,P$ **do**
    for $j = 1,\ldots,D$ **do**
      Set $rn_1 \in [0,1]$
      Set $rn_2 \in [0,1]$
$z'_{i,j} = z_{i,j} + rn_1 \cdot \left(z_{best,j} - \lvert z_{i,j} \rvert - rn_2 \cdot \left(z_{worst,j} - \lvert z_{i,j} \rvert \right) \right)$
    end for
    if $f(z'_i) \le f(z_i)$ **then**
      $z_i = z'_i$ **(Update)**
    end if
  end for
  $t = t + 1$
end while

Suppose if the objective function is $f(z)$ with dimensional D factors represented as $(j = 1,2,..,D)$ and the estimation value is $z_{i,j}$ for j^{th} variable of i^{th} competitor solution. The position of i^{th} solution candidate is $z_i = (z_{i,1}, z_{i,2}, \ldots, z_{i,D})$. The solution of best competitor can be expressed as $z_{best} = (z_{best,1}, z_{best,2}, \ldots, z_{best,D})$ has the finest estimation of $f(z)$ in the present population whereas the solution of worst competitor in the present population is $z_{worst} = (z_{worst,1}, z_{worst,2}, \ldots, z_{worst,D})$. The solution of $z_{i,j}$ is simplified using the expression,

$$z'_{i,j} = z_{i,j} + rn_1 \cdot \left(z_{best,j} - \lvert z_{i,j} \rvert - rn_2 \cdot \left(z_{worst,j} - \lvert z_{i,j} \rvert \right) \right) \tag{14}$$

Where, the values of best and worst solutions of the j^{th} variable can be represented as $z_{best,j}$ and $z_{worst,j}$. The dual random numbers are rn_1 and rn_2 within the range [0,1], the updated solution is $z'_{i,j}$ and the absolute value of $z_{i,j}$ is signified as $\lvert z_{i,j} \rvert$. In each iteration, the attraction towards the best solution is denoted as $\left(z_{best,j} - \lvert z_{i,j} \rvert \right)$. Once the finest solution is achieved via the JO algorithm, it moves towards the finest outcome neglecting the worst solution. Thus, the proposed DAE-SR model obtains top recommendations on the restaurant names by considering the user reviews. Here, DAE-SR is trained on a data. This network gains knowledge from this data, which is compiled as "weights" of the network. These weights optimized with the JO algorithm have improved the performance of DAE-SR network in recommendation system. This optimization has good exploration and exploitation ability and hence avoid local optima problem with good convergence rate. The method presented in this paper gives a new way of improving the accuracy of the system.

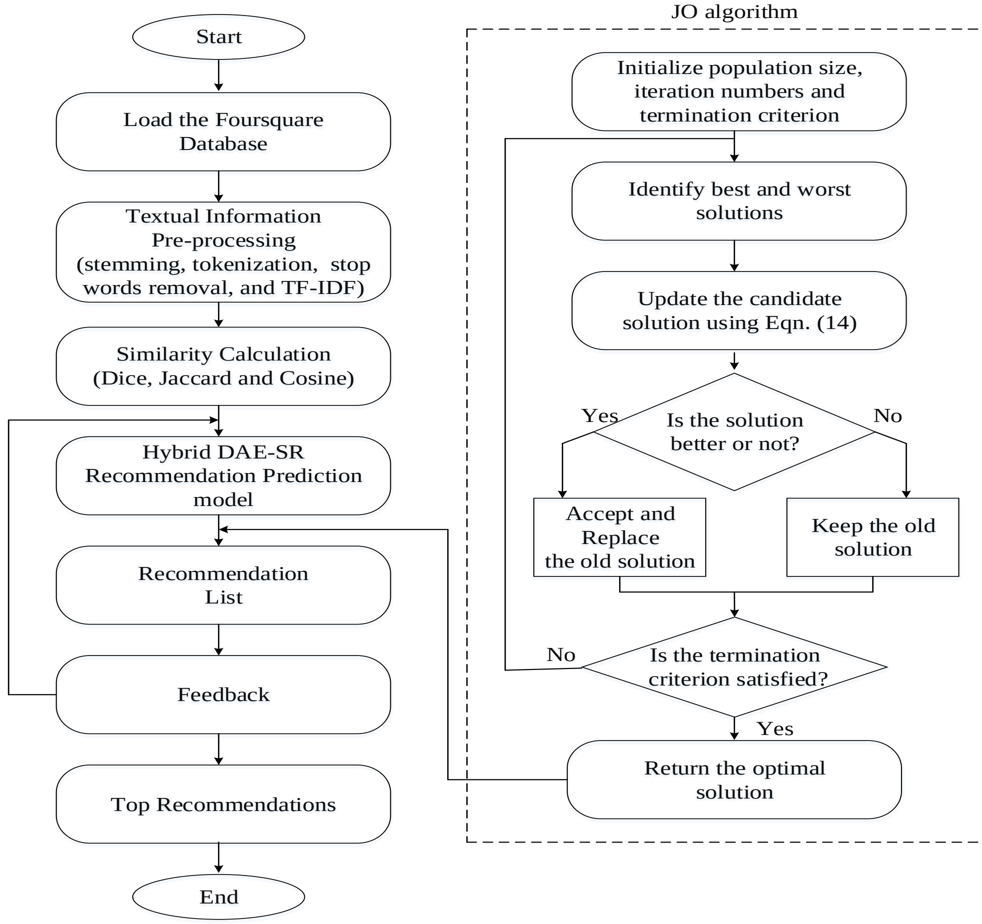

Fig. 4. Flowchart of proposed model.

Fig. 4 displays the flowchart of proposed model. The process starts by collecting the data from the Foursquare dataset. Next step is pre-processing of textual information with the stemming, tokenization, stop words removal and TF-IDF approaches. Three different similarity measures such as dice, Jaccard and cosine metrics are used to evaluate the similarity among the features. Among the three measures, cosine similarity has gained improved performance which can be used for further processing. Finally, the recommendation prediction is performed with the DAE-SR machine learning model. Here, the proposed DAE-SR model is optimized using JO algorithm that generates optimal best solutions with the feedback model and the Top recommendation on restaurant names is displayed. Some of the advantages of the proposed work are:

- Savings in cost and processing time.
- No cold-start issue and work well with sparse datasets.
- Quicker than traditional recommendation systems.
- Effective decision making and.
- Useful for people who visits new places.

E. Answers to the Research Questions

The presented context-aware recommendation model extracts valuable information from learner contexts and also leads to obtain better recommendations. In this wok, the contexts were collected using Search engine (Foursquare dataset). Next, the proposed RS named (DAE-SR) adopts a user preference model by using features, e.g., restaurant ratings, location, current weather, and time of the day. The recommendation technique used to process the context is DAE-SR with JO algorithm. The main recommended activity of the proposed model is to generate relevant and personalized recommendation about restaurants.

F. Application of the Proposed Work

Restaurant Recommendation is very useful to users because users can get information easily about different restaurants worldwide. These recommendations offer information's like popular restaurant names, menus, location, phone numbers and directions at a glance via the user's smartphone. Nowadays, there are number of social networking sites in which users post reviews about restaurants and rate businesses. However, before the creation of these online sites people went to restaurants occasionally visiting an interesting place or hearing reviews from neighbours, friends etc. With the technological advancements, there are number of online websites in which the people can check the details about the restaurant before visiting it directly. The information can be available in the form of ratings, price ranges, reviews, and hours of operation. After analysing all these factors, users can make valuable decisions. Checking the information via these online sites seem to be a time consuming task since users have to visit number of pages about different recommendations in order to find the right restaurant of their choice. Above all, when people visit a new place, travel or move to new place users prefer the same search procedure again. Therefore, the proposed DAE-SR model offers an effective recommendation system for the selection of restaurants. Using this we can build a customized real-time restaurant recommendation system. With this proposed system, we can provide recommendation for restaurants that will suit people's preference. For example, consider a newcomer who has just visited Chennai. He wants to have *'pricey but good selection of beer snobs'*. He is unsure about the good restaurants around his location. By inputting, this search query into our proposed model, he can get the top restaurants nearer to his location that are highly rated and get him interested in using our recommendation system as a user. Fig. 11 shows the best example as the application of the work in real world. Experimental results have shown that the proposed DAE-SR model achieved better performance than the existing recommendation procedures.

IV. Experimental Results

The recommender system helps the users in an effective manner to obtain useful and personalized information, in order to make comprehensive decisions in the day to day life of consumers. In experimental analysis, enactment of proposed DAE-SR is associated with the existing classifiers for example, K-Nearest Neighbour (KNN), Deep Neural Network (DNN) and Artificial Neural Network (ANN). For evaluating the proposed approach, the data is being collected from the public source and the dataset is named as Foursquare dataset. The restaurant reviews are discussed in this dataset. It contains restaurant name, location, day and time and the reviews. In terms of accuracy, precision and recall the proposed strategy is estimated. The implementation tool used in this work is PYTHON. The performance is compared based on the similarity. Dice's coefficient, Cosine similarity and Jaccard Similarity Coefficient are the tree similarity methods, which are used in the proposed strategy. Cosine similarity accuracy is better than the other similarity.

A. Description about Dataset

The dataset used for the implementation process is the Foursquare dataset. This is a publicly available data source being collected from https://www.kaggle.com/danofer/foursquare-nyc-rest?select=readme.txt. However, this dataset covers the details of check-in, tag and tip data of the restaurant locations in NYC gathered from October 24, 2011 to February 20, 2012. The foursquare dataset comprises of 3,112 users, 3,298 venues (locations) with 27,149 check-ins and 10,377 tips. Customer satisfaction can be analysed based on the reviews of reviewers acquired by the restaurants. Hence, the successful growth of restaurants is mainly based on the reviewer's reviews. It contains the details such as restaurant name, location, day and time and the reviews. Here, 70% of data is used for training and remaining 30% of data is used for testing purpose.

B. Evaluation Metrics

The procedure of analysing, collecting and reporting the data regarding the system performance is termed as performance evaluation. The performance of machine learning models is measured by means of the suitable evaluation metrics such as accuracy, recall and precision.

1) Precision value: It is the ratio of total number of recommendations that are relevant among the number of total recommendations provided.

$$Precision = \frac{TP}{TP + FP} \tag{15}$$

2) Recall value: It is the ratio of number of recommendations provided among the total relevant recommendations.

$$Recall = \frac{TP}{TP + FN} \tag{16}$$

3) Accuracy value: This metric gives the required related documents from the total documents. The accuracy metric tries to evaluate the effective decision making of recommendation algorithms. It can be used to estimate the amount of correct as well as incorrect classifications as relevant or irrelevant items which can be predicted by the recommender system and hence useful for user tasks such as finding good items.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{17}$$

(TP- True positive, TN- True Negative, FP- False Positive, FN- False Negative).

C. Performance Analysis

The analysis of proposed DAE-SR is shown and compared with DNN, ANN, KNN for Dice's Coefficient, Jaccard similarity and cosine similarity are explained in this section.

Table VII gives the performance value of DNN, ANN, KNN and proposed DAE-SR for Dice's Coefficient. Fig. 5 displays the comparison analysis of dice similarity coefficient. For DNN, ANN, KNN and proposed DAE-SR classifier the precision values obtained are 85.4, 88.5, 91 and 95.16%, followed by recall value 80.2, 84.9, 87.9 and 94.22% respectively. The accuracy values for DNN, ANN, KNN and proposed DAE-SR classifier are 85.3, 90, 90.4, and 94.66%. Compared with the existing techniques our proposed approach gained better performance.

TABLE VII. DICE'S COEFFICIENT BASED PERFORMANCE

Performance Metrics	KNN (%)	ANN (%)	DNN (%)	DAE-SR (Proposed) (%)
Precision	85.4	88.5	91	95.16
Recall	80.2	84.9	87.9	94.22
Accuracy	85.3	90	90.4	94.66

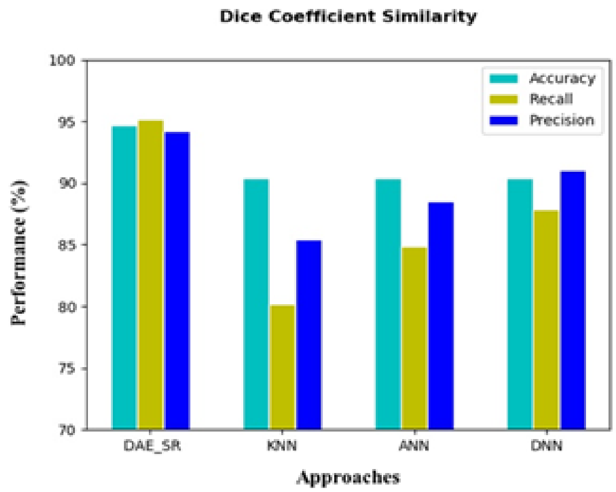

Fig. 5. Performance analysis of DNN, ANN, KNN and proposed DAE-SR for dice's coefficient.

TABLE VIII. PERFORMANCE ANALYSIS VALUE OF JACCARD SIMILARITY

Performance Parameter	KNN	ANN	DNN	DAE-SR (Proposed)
Precision	87.9	90	92.3	97
Recall	83.4	86.2	89	96.7
Accuracy	86.7	91.5	93	97.25

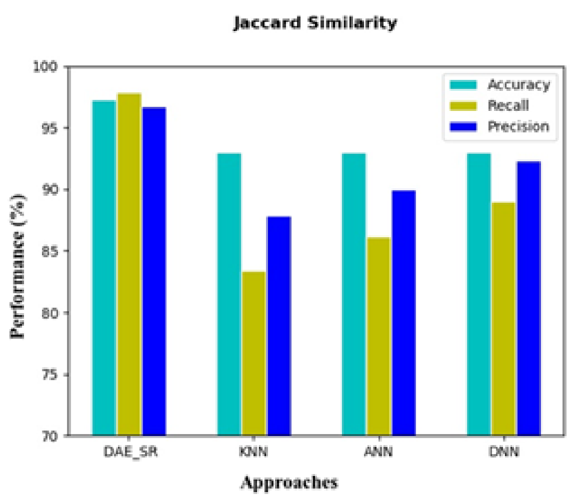

Fig. 6. Graphical representation of jaccard similarity performance.

Table VIII represents the Jaccard Similarity performance value comparison of different classifiers with proposed classifier. Fig. 6 shows the analysis of DNN, KNN, ANN and proposed DAE-SR for Jaccard similarity. For DNN, ANN, KNN and proposed DAE-SR classifier precision value are 87.9, 90, 92.3 and 97%, recall value are 83.4, 86.2, 89 and 96.7% also accuracy value are 86.7, 91.5, 93 and 97.25%. It is obviously agreed, compared with other classifier the proposed approach has better performance.

TABLE IX. PERFORMANCE OF COSINE SIMILARITY

Performance Metrics	KNN	ANN	DNN	DAE-SR (Proposed)
Precision	90	92	94	98
Recall	85	88	92	98.1
Accuracy	89	93	96	98.33

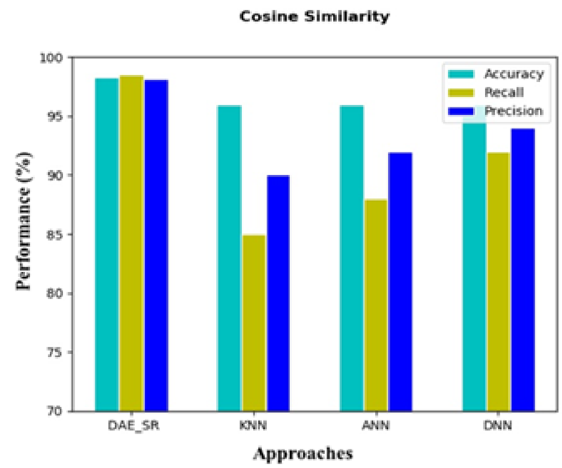

Fig. 7. Graphical representation of DNN, ANN, KNN and proposed DAE-SR for cosine similarity performance.

Table IX signifies the performance of Cosine similarity. Fig. 7 displays the analysis of DNN, ANN, KNN and proposed DAE-SR classifier on cosine similarity. The accuracy acquired with DNN, ANN, KNN and DAE-SR is 89, 93, 96 and 98.33% respectively. From the above results, it is finely known compared with other classifier the proposed DAE-SR approach achieved better performance. Also, cosine similarity based performance is high when compared to the other similarity. As in our proposed method, the generating value accuracy is better in Cosine similarity. Table X represents the comparative results of RMSE and MAE. The proposed DAE-SR model shows lower error values which clearly signify the accurateness of the system compared to the existing methods.

Fig. 8 displays the comparison outcomes of RMSE and MAE of the proposed approach with existing approaches. Minimum or decreased in RMSE and MAE is considered as better outcome. Error minimizing gives better accuracy for the proposed scheme compared to other approaches. In figure 8, compared to two existing approaches, the proposed strategy produced minimum error outcomes.

TABLE X. COMPARATIVE RESULTS OF RMSE AND MAE

Approaches	RMSE	MAE
DAE-SR (Proposed)	0.15	0.02
Natarajan *et al.* (2020) [25]	0.9	0.8
Iqbal *et al.* (2019) [26]	0.8832	1.092

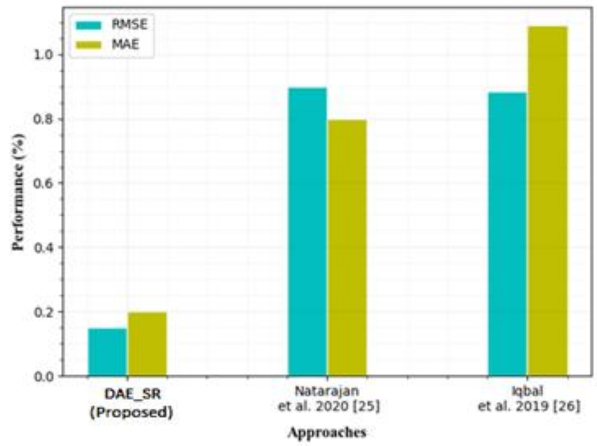

Fig. 8. Comparison of different performance measures.

TABLE XI. COMPARISON OF PERFORMANCE OUTCOMES OVER DIFFERENT EXISTING WORKS

Method/Metric	Precision	Recall	Accuracy
DAE-SR (proposed)	98%	98.1%	98.33%
Natarajan *et al.* (2020) [25]	56.4%	56%	84.1%
Afolabi *et. al.* (2019) [22]	46.9%	42.2%	-

Table XI gives the various performance outcomes comparison regarding different existing works. Fig. 9 shows the precision, recall, and accuracy comparison of DAR-SR with different existing works. In this research, we have used DAE-SR for the effective recommended system. The proposed approach produced better outcomes than the other two previous works. Our proposed model achieves maximum 98% precision, 98.1% recall and 98.33% accuracy values. The proposed scheme achieves above 98% performances because of the best classification approach DAE-SR with JO.

TABLE XII. ACCURACY COMPARISON WITH EXISTING METHODS

Authors	Techniques	Accuracy (%)
Venugopal and Sandhya	**DAE-SR (proposed)**	98.33%
Natarajan *et al.*[25]	RS-LOD	84.1%
Ray *et al.*[34]	Ensemble	92.36%
	RNN	86.00%
	GRU	90.00%
	LSTM	89.00%
	Bi-LSTM	89.00%

Table XII signifies the accuracy comparison with existing methods. The accuracy of the proposed DAE-SR model is compared with various state-of-the-art models to evaluate the performance of the proposed recommendation procedure. The existing models on recommendation system gained lower classification accuracy with RS-LOD, RNN, Ensemble, LSTM, GRU and Bi-LSTM as 84.1%, 86.00%, 92.36%, 89.00%, 90.00% and 89.00% respectively. The proposed model achieved supreme accuracy of (98.33%) because the context-aware RS based on text mining and machine learning algorithms afford quick and correct suggestions to user on good restaurants name. Table XIII specifies the run time comparison with existing methods. From the table it is observed that the proposed DAE-SR recommendation model requires less run time 1.43 milliseconds for computation whereas the other models specify higher computational time which reduces the system efficacy.

TABLE XIII. RUN TIME COMPARISON WITH EXISTING METHODS

Techniques	Run Time
DAE-SR (proposed)	1.43 ms
IDA-CF [35]	2.6 ms
CF [35]	12 sec
NRA [35]	27 sec
FRS [35]	28 sec

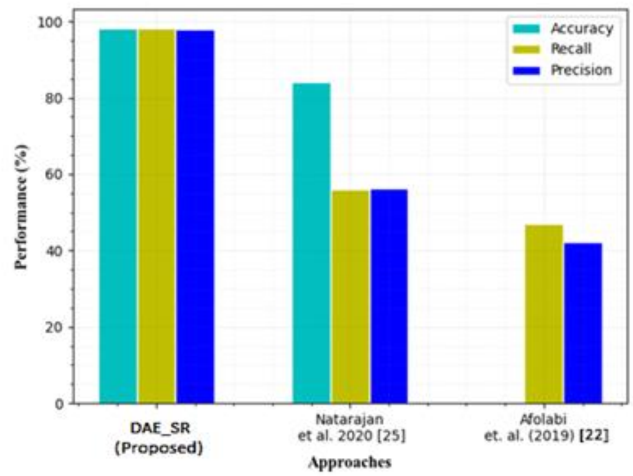

Fig. 9. Performance comparison of DAE-SR.

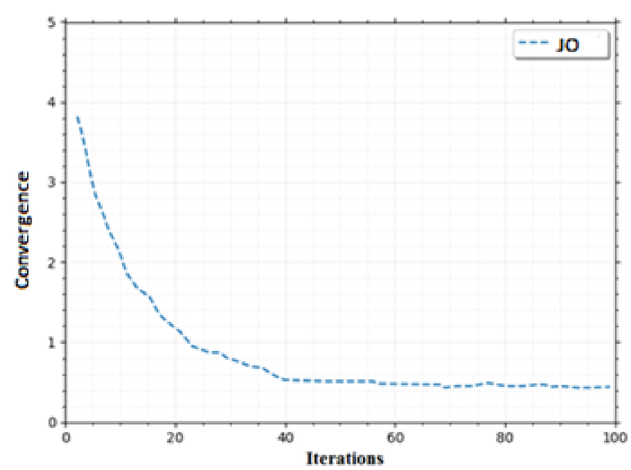

Fig. 10. Convergence graph for JO.

Fig. 10 demonstrated the convergence graph for the JO algorithm. It illustrated that, the proposed JO algorithm achieved the optimum convergence after the 40 iterations, because the JO is a potential optimizer for an engineering issues. Compared to other algorithms, the exploration and exploitation ability of JO algorithm is very impressive and superior. And it has a very good balance among exploitation and exploration.

D. Recommendation Analysis

The following section gives the recommendation analysis in terms of a graphical user interface (GUI). Similarity-based outcomes are taken here for the statistical measures and compared with the existing classifiers. The proposed model DAE-SR achieved better performance, and it produced accurate recommendation outcomes.

Fig. 11. Content based restaurant recommendation.

Recommendation of the good restaurants name, location and the reviews of the recommender is displayed in this analysis. Fig. 11 shows the content based restaurant recommendation. The recommendation text box shows the restaurant name, location, day and time also the customer reviews. There are large amount of data is there. The user's information that is related to the restaurant is shown by GUI.

E. Discussion

The proposed study introduces a DL based restaurant recommendation system named DAE-SR which extracts user preferences from online comments and recommends top restaurants to the users. With the technological advances and changes in human lifestyle makes the selection of appropriate restaurants as very difficult due to the wide range of ethnicities, ingredients, culinary styles, personal tastes and cultures. The proposed DAE-SAR based recommendation systems offers valuable decisions to users on where to eat. The proposed DAE-SR model attained an increased accuracy of 98.33% whereas the existing models RNN, LSTM, Ensemble, GRU, BiLSTM and RS-LOD gained lower results such as 86%, 89%, 92.36%, 90%, 89% and 84.1% respectively. Thus, the presented model obtained improved accuracy due to better convergence, low run-time and GUI modelling.

V. CONCLUSION

Recommender system provides precise online recommendations to the web users, it can also be known as knowledge management and information distribution tool. In this paper, a novel distributed framework for context aware recommender system is proposed based on text mining and machine learning algorithms. Based on the user opinion feedback reviews from the social networking services the top restaurants are recommended to the customers. The existing works on recommendation system proved that the use of individual CF based recommendation provide recommendations only via ratings/user interests and items. In this work, a new RS model is presented which proved that the contextual information is more relevant in the generation of recommendations under diverse conditions such as ubiquity and mobility. The proposed DAE-SR classifier offers suitable suggestions to users on the restaurant recommendations compared to the existing DNN, KNN and ANN methods. Experimental outcomes reveal that the proposed DAE-SR based recommendation system outperforms the other classifiers in terms of accuracy, precision, and recall values based on the similarity measures. The accuracy (98.33%) of proposed strategy DAE-SR outperforms the other strategies with cosine similarity. The main advantage of DAE-SR is, it may not suffer from cold-start issue and work well with sparse datasets. The limitation of this study is that only one dataset is used for analysis. Moreover, it is limited by the total amount of available content for providing meaningful recommendation. In future work, we can evaluate our proposed method with different datasets for different applications.

REFERENCES

[1] C. Rodney, "Modern communication technologies and the new world information order". International journal of communication vol. 2, no. 1, pp. 147-159, 2005.

[2] J. Ahokas and T. Kiiski, "Cybersecurity in ports". Publication of the HAZARD Project vol. 3, 2017.

[3] F. Iannelli, A. Koher, D. Brockmann, P. Hövel and I.M. Sokolov, "Effective distances for epidemics spreading on complex networks". Physical Review E vol. 95, no. 1, pp. 012313, 2017.

[4] R. Pastor-Satorras, C. Castellano, P.V. Mieghem and A. Vespignani, "Rev. Mod. Phys" vol. 87, pp. 925, 2015.

[5] A.S. Manek, P.D. Shenoy, M.C. Mohan and K.R. Venugopal, "Aspect term extraction for sentiment analysis in large movie reviews using Gini Index feature selection method and SVM classifier". World Wide Web vol. 20, no. 2, pp. 135-154, 2017.

[6] G. Scavo, Z.B. Houidi, S. Traverso, R. Teixeira and M. Mellia, "WeBrowse: Leveraging User Clicks for Content Discovery in Communities of a Place". Proceedings of the ACM on Human-Computer Interaction CSCW vol. 1, no. 93, 2017.

[7] D.L. Quoc, C. Fetzer, P. Felber, "UniCrawl: a practical geographically distributedweb crawler". In: IEEE International Conference on Cloud Computing, IEEE, pp. 389–396, 2015.

[8] Y. Tang, L. Wei, W. Wang and P. Xuan, "Towards Intelligent Web Crawling–A Theme Weight and Bayesian Page Rank Based Approach". In International Conference on Web Information Systems Engineering Springer, Cham pp. 192-201, 2017.

[9] A. Darer, O. Farnan and J. Wright, "Automated Discovery of Internet Censorship by Web Crawling" 2018. arXiv preprint arXiv: 1804.03056.

[10] P. Bandagale, N.R. Sawantdesai, R.U. Paradkar and P.P. Shirodkar, "Survey on Effective Web Crawling Techniques". International Journal of Computer & Mathematical Sciences IJCMS, vol. 6, no. 10, October 2017.

[11] S.P. Venugopal, "Focused crawling from the basic approach to context aware notification architecture". Indonesian Journal of Electrical Engineering and Computer Science vol. 13, no. 2, pp. 492-498, 2019 10.11591/ijeecs.v13.i2.

[12] F. Xiao, C. Zhan, H. Lai, L. Tao and Z. Qu, "New parallel processing strategies in complex event processing systems with data streams". International Journal of Distributed Sensor Networks vol. 13, no. 8, pp. 1550147717728626, 2017.

[13] G. Cugola and A. Margara, “Processing flows of information: from data stream to complex event processing”. ACM ComputSurv vol. 44, no. 3, pp. 15, 2012.

[14] O.J. Lee and J.E. Jung, “Sequence clustering-based automated rule generation for adaptive complex event processing”. Future Generation Computer Systems vol. 66, pp. 100-109, 2017.

[15] K. Patroumpas, T.K. Sellis, “Event processing and real-time monitoring over streaming traffic data, in: S.D. Martino, A. Peron, T. Tezuka (Eds.)”, Proceedings of the 11th International Symposium on Web and Wireless Geographical Information Systems, W2GIS 2012, in: Lecture Notes in Computer Science, Springer, Naples, Italy vol. 7236, pp. 116–133, 2012.

[16] I. Flouris, N. Giatrakos, A. Deligiannakis, M. Garofalakis, M. Kamp and M. Mock, “Issues in complex event processing: Status and prospects in the big data era”. Journal of Systems and Software vol. 127, pp. 217-236, 2017.

[17] P. Gopalan, L. Charlin and D.M. Blei, “Content-based recommendations with poisson factorization”. In Proceedings of NIPS, pp. 3176–3184, 2014.

[18] J. He, H.H. Zhuo and J. Law, “Distributed-Representation Based Hybrid Recommender System with Short Item Descriptions” 2017. arXiv preprint arXiv: 1703.04854.

[19] C. Li, G. Chen and F. Wang, “Recommender systems based on user reviews: the state of the art”. User Modeling and User-Adapted Interaction vol. 25, no. 2, pp. 99-154, 2015.

[20] P. Bhattacharjee and A. Awekar, “Batch Incremental Shared Nearest Neighbour Density Based Clustering Algorithm for Dynamic Datasets”. In European Conference on Information Retrieval, Springer, Cham, pp. 568-574, 2017.

[21] Y. Xia, G.D. Fabbrizio, S. Vaibhav and A. Datta, “A Content-based Recommender System for E-commerce O ers and Coupons” 2017.

[22] I.T. Afolabi, O.S. Makinde and O.O. Oladipupo, “Semantic Web mining for Content-Based Online Shopping Recommender Systems”. International Journal of Intelligent Information Technologies (IJIIT) vol. 15, no. 4, pp. 41-56, 2019.

[23] F. Amato, V. Moscato, A. Picariello and F. Piccialli, “SOS: A multimedia recommender System for Online Social networks”. Future generation computer systems, Springer 2017.

[24] J.K. Tarus, Z. Niu and A. Yousif, “A hybrid knowledge-based recommender system for e-learning based on ontology and sequential pattern mining”. Future Generation Computer Systems vol. 72, pp. 37-48, 2017.

[25] S. Natarajan, S. Vairavasundaram, S. Natarajan and A.H. Gandomi, “Resolving data sparsity and cold start problem in collaborative filtering recommender system using Linked Open Data”. Expert Systems with Applications vol. 149, pp. 113248, 2020.

[26] M. Iqbal, M.A. Ghazanfar, A. Sattar, M. Maqsood, S. Khan, I. Mehmood and S.W. Baik, “Kernel context recommender system (KCR): A scalable context-aware recommender system algorithm”. IEEE Access vol. 7, pp. 24719-24737, 2019.

[27] A. Pujahari and D.S. Sisodia, “Pair-wise Preference Relation based Probabilistic Matrix Factorization for Collaborative Filtering in Recommender System”. Knowledge-Based Systems pp. 105798, 2020.

[28] M.H. Aghdam, “Context-aware recommender systems using hierarchical hidden Markov model”. Physica A: Statistical Mechanics and its Applications vol. 518, pp. 89-98, 2019.

[29] H.M. Pandey, “Jaya a novel optimization algorithm: What, how and why?” In 2016 6th International Conference-Cloud System and Big Data Engineering (Confluence) IEEE pp. 728-730, 2016.

[30] A.K. Sangaiah, D.V. Medhane, T. Han, M.S. Hossain and G. Muhammad, “Enforcing position-based confidentiality with machine learning paradigm through mobile edge computing in real-time industrial informatics”. IEEE Transactions on Industrial Informatics vol. 15, no. 7, pp. 4189-4196, 2019.

[31] A.K. Sangaiah, D.V. Medhane, G.B. Bian, A. Ghoneim, M. Alrashoud and M.S. Hossain, “Energy-aware green adversary model for cyberphysical security in industrial system”. IEEE Transactions on Industrial Informatics vol. 16, no. 5, pp. 3322-3329, 2019.

[32] A.K. Sangaiah, A.A.R. Hosseinabadi, M.R. Shareh, S.Y.B. Rad, A. Zolfagharian and N. Chilamkurti, “IoT resource allocation and optimization based on heuristic algorithm”. Sensors vol. 20, no. 2, pp. 539, 2020.

[33] A.K. Sangaiah, M. Sadeghilalimi, A.A.R. Hosseinabadi and W. Zhang, “Energy consumption in point-coverage wireless sensor networks via bat algorithm”. IEEE Access vol. 7, pp. 180258-180269, 2019.

[34] B. Ray, A. Garain, and R. Sarkar, “An ensemble-based hotel recommender system using sentiment analysis and aspect categorization of hotel reviews”. Applied Soft Computing vol. 98, pp. 106935, 2021.

[35] B. Ramzan, I.S. Bajwa, N. Jamil and F. Mirza, “An Intelligent Data Analysis for Hotel Recommendation Systems using Machine Learning” 2019. arXiv preprint arXiv:1910.06669.

[36] G. Gupta and R. Katarya, “EnPSO: An AutoML Technique for Generating Ensemble Recommender System." Arabian Journal for Science and Engineering pp. 1-19, 2021.

[37] R. Katarya and Y. Arora, “Capsmf: a novel product recommender system using deep learning based text analysis model”. Multimedia Tools and Applications vol. 79, no. 47, pp. 35927-35948, 2020.

[38] R. Katarya, “Reliable Recommender System Using Improved Collaborative Filtering Technique”. System Reliability Management: Solutions and Technologies vol. 113, 2018.

[39] R. Katarya, “Movie recommender system with metaheuristic artificial bee”. Neural Computing and Applications vol. 30, no. 6, pp. 1983-1990, 2018.

[40] D.K. Yadav and R. Katarya, “Study on Recommender System using Fuzzy Logic”. In 2018 Second International Conference on Computing Methodologies and Communication (ICCMC), IEEE, pp. 50-54, 2018.

[41] R. Katarya, “An improved recommender system using two-level matrix factorization for product ontologies”. In 2017 International Conference on Intelligent Sustainable Systems (ICISS), IEEE, pp. 223-226, 2017.

[42] R. Katarya and O.P. Verma, “An effective web page recommender system with fuzzy c-mean clustering”. Multimedia Tools and Applications vol. 76, no. 20, pp. 21481-21496, 2017.

[43] R. Katarya and O.P. Verma, “Efficient music recommender system using context graph and particle swarm”. Multimedia Tools and Applications vol. 77, no. 2, pp. 2673-2687, 2018.

[44] R. Katarya, M. Ranjan and O.P. Verma, “Location based recommender system using enhanced random walk model”. In 2016 Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC), IEEE, pp. 33-37, 2016.

[45] R. Katarya, O.P. Verma and I. Jain, “User behaviour analysis in context-aware recommender system using hybrid filtering approach”. In 2013 4th International Conference on Computer and Communication Technology (ICCCT), IEEE, pp. 222-227, 2013.

[46] C.N. Sánchez, J. Domínguez-Soberanes, A. Arreola and M. Graff, “Recommendation System for a Delivery Food Application Based on Number of Orders”. Applied Sciences vol. 13, no. 4, pp. 2299, 2023.

[47] B. Teixeira, D. Martinho, P. Novais, J. Corchado and G. Marreiros, “Diabetic-Friendly Multi-agent Recommendation System for Restaurants Based on Social Media Sentiment Analysis and Multi-criteria Decision Making.” In Progress in Artificial Intelligence: 21st EPIA Conference on Artificial Intelligence, EPIA 2022, Lisbon, Portugal, August 31–September 2, 2022, Proceedings, Cham: Springer International Publishing, pp. 361-373, 2022.