



Hybrid Approach to Recommender System Model

Abhilasha Sankari^{1*}, Shraddha Masih², Maya Ingle³

^{1*} School of Engineering, Renaissance University, Indore

^{2,3} School of Computer Science & Information Technology, DAVV University Indore

*abhilasha.renaissance@gmail.com

Abstract. Recommender systems are a popular research area broadly applied, from e-commerce to e-learning systems. The paper presents a hybrid approach to recommender systems that leverages user reviews, clustering, and sentiment analysis to enhance recommendation accuracy. Traditional models primarily rely on user ratings, often neglecting the rich contextual information in textual reviews. The proposed method extracts user preferences by analyzing re-view content through sentiment analysis and clustering techniques, providing a more nuanced understanding of user tastes and preferences. By incorporating this detailed preference data and contextual information, the system generates Top-N recommendations that are relevant to the user and personalized. We applied our model to the Yelp dataset, which includes diverse and extensive user reviews of various businesses. Comparative evaluations demonstrate that our approach significantly outperforms traditional models based solely on user ratings, achieving higher accuracy in recommendation predictions. The results underscore the importance of utilizing multidimensional data sources in recommendation systems, highlighting the potential for improved user satisfaction and engagement. This study contributes to recommender systems by showcasing the benefits of a comprehensive analysis of user-generated content and its impact on recommendation quality.

Keywords: Recommender Systems, Hybrid Recommender System, Collaborative Filtering

1 Introduction

Nowadays, recommendations have become a part of everyday life. Recommender systems techniques have been sturdily researched within the last decade to facilitate the demand for prioritizing information and suggesting information of potential interest to the users. The recommender system is widely used in various domains, including resumes, movie recommendations, e-commerce, etc. It has become an effective way in marketing processes in e-commerce because it provides valuable and personalized information. Our paper aims to develop a hybrid approach to recommendation that can exploit implicit feedback and sentimental analysis for calculating user preferences and suggesting appropriate items according to user preferences.

© The Author(s) 2025

S. Bhalerao et al. (eds.), *Proceedings of the International Conference on Recent Advancement and Modernization in Sustainable Intelligent Technologies & Applications (RAMSITA-2025)*, Advances in Intelligent Systems Research 192,

https://doi.org/10.2991/978-94-6463-716-8_28

2 Related Work

There are two approaches to recommender systems: the classical 2D approach and the multidimensional or contextual approach [1]. The classical approach or two-dimensional approaches to recommender systems only consider the rating matrix of user and item. It emphasizes recommending the most appropriate items to users without incorporating any contextual information. In many applications, such as suggesting a restaurant, it is inadequate to consider only a combination of users and items. Incorporating contextual information into the recommendation process is vital for providing users with item suggestions tailored to specific conditions. Reviews crawled from the people who had already visited the restaurant are valuable sources of information for recommendation. Therefore, accurate user preference prediction depends on how well the appropriate contextual information is combined into a recommendation process.

3 Context Aware Recommender System (CARS)

Context is any aspect or sub-aspect of an entity that can be used to characterize a situation. In [2], context is “Any information that can be used to characterize the situation of an entity. An entity is a person, place, or object considered relevant to the interaction between a user and an application, including the user and applications themselves”. The application incorporates the user's contextual information and adapt dynamically to changes in this context to deliver precise recommendations. This is known as context awareness [3] [4]. Table 1 shows the example of context for recommendation.

Table 1. Example of Context for Recommendation

S.No.	Reference	Context/Aspect
1	[2]	Context toolkit to build context-aware applications
2	[5]	Identity, location, objects, nearby people
3	[6]	Season, location, identity, object, nearby people
4	[7]	Location, identity, time, activity, Email, Phone Number, Address, Date of Birth
5	[8]	feature of interaction/emergent property of occasions of interaction/ characteristics of the environment in which the activity occurs
6	[9]	Where you are/ Who you are with/What resources are nearby
7	[10]	State, reachability, surroundings, location
8	[11]	Location, time, weather, user preferences

Context Information taken from user opinion increases the accuracy of recommendation systems. In the context of restaurant recommendations, the factors to consider include the food quality, service type, location, and the ambiance of the venue. For example, "food quality" can be a vital feature for some users, while saving traveling time is essential for others. Location is the most crucial feature in restaurant recommendation systems among the many contextual information [12]. To accurately

predict the preferences and provide better recommendations, the Context Aware Recommender System incorporates context information into the classical recommender system [13] [14].

4 Sentiment Analysis Methods for Recommender System

Sentiment analysis is a computational technique to study the feelings, emotions and user attitude towards any individual, event, place, topic, service, product, and their attributes [15]. It is also called opinion mining. It is an interesting part of machine learning. The focus of sentiment analysis is to extract polarities from data. Over the past five years, many surveys [16, 17, 18] have been performed on sentiment analysis methods for recommender systems. These are good starting points for discovering new challenges in recommendations. Complexity is high in natural language processing since the same word can express different things. In such cases, a semantic approach to sentiment analysis is suitable.

Carl Anderson in [19] presents a literature survey on technical approaches to food recommendations, such as collaborative filtering, content-based recommendations, hybrid, graph-based recommendations, word embeddings, expert systems, and clustering. Hegde et al. in [20] utilized general approaches like Bag of Word and Term frequency-inverse document frequency (Tf-idf) to extract the features of the food items. Utilizing the Support Vector Machine algorithms, user feedback is analyzed based on extracted features, and top-n food items are recommended to be included. Table 2 shows the summary of work related to sentiment analysis methods for recommendation. Table 3 shows example of restaurant reviews consisting of features and sentiments.

Table 2. Summary of Work Related to Sentiment Analysis Methods for Recommendation

Year, author, reference	Main Subject	Dataset	Proposed Method	Conclusion	Future Work
[21]	Comparative study of Rule-Based Sentiment Analysis, Aspect-Based Sentiment Analysis, and Sentiment Analysis using Machine Learning	Kaggle Dataset, Movie Review, aclimbstandford, Rt-polarity dataset	Aspect-based Sentiment analysis is integrated with maximum entropy.	Machine Learning has higher accuracy.	A parallelized ranking algorithm can be incorporated to improve results.
[22]	Multi-criteria recommendation using aspect-based sentiment analysis	Yelp, Tripadvisor, Amazon	Framework for Sentiment Aspect Based Retrieval (SABRE)	Framework extracts relevant aspects and their sub-aspects from reviews	Encode the extracted features from users' reviews on the context aware recommendation algorithm. Extract negative sentiment from customer review
[20]	Sentiment Classification and four-course meal classification	Yelp	Framework for sentiment-based food classification	Address the issue in identifying top-n food items of users' preference	ARP can be applied on multiple datasets to study the effects of sparsity and dimensionality features over its performance
[23]	Attractive-Relevance-Popularity measures performance compared to traditional similarity measures in a collaborative filtering context.	Yelp Restaurants Reviews/ Datafiniti Hotel Review	Proposed sentiment rating approach using SentiwordNet /Designed Attractive-Relevance-Popularity measure to determine the similarity between two users.	ARP successfully replaces traditional similarity measure like Pearson Correlation Coefficient.	

Table 3. Example of Restaurant Reviews Consisting of Features and Sentiments

1.	The restaurant was decent. The food was okay, but nothing really stood out. I ordered the burger, and while it was cooked well, the toppings were a bit lacking.
2.	Absolutely love this place! The selection of craft beers is impressive, and the staff is incredibly knowledgeable.
3.	I was really disappointed with my visit. The wait time was excessively long, and when our food finally arrived, it was cold.
4.	The staff was friendly, but it took a while for our food to arrive. I might come back to try something else.
5.	This place receives stars for their APPETIZERS!!!

5 Problem Statement

Most traditional restaurant recommendation system research is based on the user's rating history. In past research, user opinions are seldom considered. Therefore, the accuracy of the recommendation is not satisfactory. For example, if user A's opinion about a restaurant is negative, then user B, who has the same taste as user A, should not recommend that restaurant. Restaurant recommendations are quite difficult, as users hardly rate restaurants without bias. Due to this, it is challenging to build a user profile for recommendations. This lack of information or no information at all leads to a cold start problem [24]. Therefore, rather than focusing on ratings and users' profiles, the proposed model takes users' reviews and generates user preferences accordingly. This method is more effective in situations where ratings are unavailable.

6 Proposed Model

The proposed model (Fig 1) comprises three phases: extracting the user preferences, incorporating contextual information, and making recommendations.

6.1 Extracting preferences

In this phase, the following steps are preprocessing, clustering, and sentiment analysis to extract the user preference. Information filtering techniques are used to find relevant information in a dynamic database. The first step is preprocessing, which includes tokenization, POS tagging, stop word elimination, stemming, and extracting nouns.

Tokenization is the process of splitting text into individual words or tokens. The basic phase in NLP involves splitting a text into meaningful units. In POS tagging, a part-of-speech tag is assigned to each token, identifying its grammatical category, such as noun, verb, or adjective. Here, we filter out only the nouns. The stop word elimination method removes common stop words (words that don't contribute much to the meaning) from the text. Nltk provides a predefined list of stop words.

Stemming is utilized to convert the words to their base or root form. This helps in dropping the inflected words to a common base form.

After noun extraction from the user's opinion, nouns that are not related to the restaurant domain are identified. In the next step, the semantic similarity of the related nouns is identified to form a cluster. We used Word embedding to identify semantic similarity among related nouns and form clusters. Word Embeddings represent words in a continuous vector space. In our experiment, we utilized Word2Vec word embedding models. It maps words into high-dimensional vectors, capturing semantic relationships. A similarity graph based on the semantic similarity of the extracted nouns is created, utilizing the cosine similarity matrix [25]. The next step is clustering. Once the clustering is done, we applied sentimental analysis to explore the information and the textual reviews that are relevant for recommendation. The clusters are scored using sentiment analysis. The user preferences are identified through the cluster with the highest sentiment score.

6.2 Algorithm for User Preference

Inputs: User Reviews

Contextual Information

Output: Top-N list of User Preferences

Step 1: Preprocessing

- i. Tokenization
- ii. POS Tagging
- iii. Stop-Word Elimination
- iv. Stemming
- v. Extracting Noun

Step 2: Clustering

Step 3: Sentiment Analysis

- i. Sentiment feature extraction and selection
- ii. Sentiment Classification (positive, neutral, negative)

Step 4: Score each cluster based on sentiment analysis and noun frequency

Step 5: Select the cluster having the highest score

6.3 Incorporating context or side information

The use of context or side information such as geographic locations with user preference could favor the recommendation process [26]. Some contextual information in the restaurant recommender system is the weather, location, time, and business attributes such as "Good For Kids," "Has TV," etc. The aim of incorporating context information in the recommender system is to improve user satisfaction and also the accuracy of the recommender system.

6.4 Recommendation

A personalized restaurant recommender system model provides the Top-N recommendation list to the target user by combining user preferences and contextual information. The user preferences are then converted into vectors. The preference vector is matched with the contextual information vector using cosine similarity to find the similarity between them. The Top-N restaurant most similar to the user's preferences is recommended based on his context.

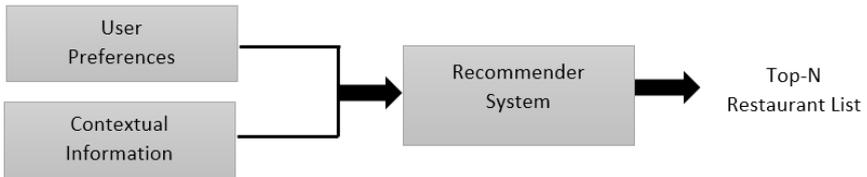


Fig. 1. CARM Model

7 Results and Discussion

The proposed model is evaluated using the Yelp Open Dataset(<https://www.yelp.com/dataset>), which comprises 6,990,280 reviews of 150,346 businesses in 11 metropolitan areas. The dataset consists of review data, business data, user data, check-in data, tip data, and photo data. We limited our experiment to the restaurant business in the dataset and omitted check-in, tip, and photo data. The traditional restaurant recommender systems utilized only rating information. The paper compares the proposed model against the traditional rating-based restaurant recommender system.

Precision, recall, and F-measure are essential metrics for assessing the performance of recommender systems. Precision focuses on the percentage of recommended items that are actually useful to the user. It is calculated by dividing the number of relevant items found by the total number of items suggested. This metric reflects the accuracy of the recommendation system. Recall evaluates how many relevant items the system successfully recommends from the entire set of relevant items. It is determined by dividing the number of relevant items retrieved by the total number of relevant items available. This metric shows the system's effectiveness in finding all possible relevant items. F-measure is the harmonic mean of precision and recall, balancing the two. It is instrumental when the dataset has an uneven class distribution, ensuring that precision and recall are considered in the evaluation. Table 4 indicates the performance evaluation of CARM Model on Yelp Dataset using precision, recall and F1-measure

Table 4. Performance Evaluation of CARM Model on Yelp Dataset using Precision, Recall and F1-Measure

	Top 5 Recommendation			Top 10 Recommendation		
	Precision	Recall	F-Measure	Precision	Recall	F-Measure
Rating-Based Model	11.67	4.22	0.051	11.88	7.30	0.072
Proposed Model	18.51	5.89	0.073	19.50	8.91	0.099

The evaluation uses Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). These are metrics used to evaluate the accuracy of predictive models. MAE measures the average magnitude of errors. The Mean Absolute Error (MAE) is calculated by taking the average of the absolute differences between predicted and actual values, giving a simple measure of the overall error. On the other hand, the Root Mean Square Error (RMSE) measures the average error size but places more emphasis on larger errors. RMSE is determined by taking the square root of the average of the squared differences between predicted and actual values. This metric is useful for identifying significant errors and is more responsive to outliers compared to MAE. Table 5 shows the performance evaluation of traditional model of CARM on Yelp Dataset using MAE and RMSE.

We have used Top-5 and Top-10 recommendations to evaluate our model. In the Top-5 recommendations, 11.67% of the suggestions made by the rating-based model are relevant, whereas the CARM model achieves 18.51% relevant viz. much higher than the rating-based model. The rating-based model captures 4.22% of all relevant items and compared with 5.89% for the CARM model, showing a slight improvement (Fig 1). The F-Measure score of 0.051 shows that the model performs low when balancing both precision and recall for Top-5 (Fig. 2) recommendations for rating-based model. The F-Measure of 0.073 indicates better overall performance compared to rating-based model when considering both precision and recall for CARM model. Context Aware Recommendation Model demonstrates superior performance across all metrics, suggesting that it delivers more accurate and relevant recommendations within the Top-5 while still identifying more relevant items.

In rating-based model, 11.88% of Top-10 recommendations are relevant, while 19.50% of Top-N recommendations are relevant for CARM. This indicates its improvement over both the rating-based model and its own Top-5 performance. In rating-based model for Top-10 recommendations, 7.30% of relevant items are successfully captured, a notable improvement from the 4.22% recall seen with the Top-5 (Fig 3). In CARM, for the Top-10 recommendations, 8.91% of all relevant items are captured, indicating a better ability to identify relevant items compared to the rating-based model.

The F-Measure of 0.072% for Top-10 recommendations in rating-based model indicates that while recall is improved, the overall balance between precision and recall is still low. The F-Measure of 0.099 reflects a stronger overall performance, with the CARM model effectively balancing precision and recall (Fig 4).

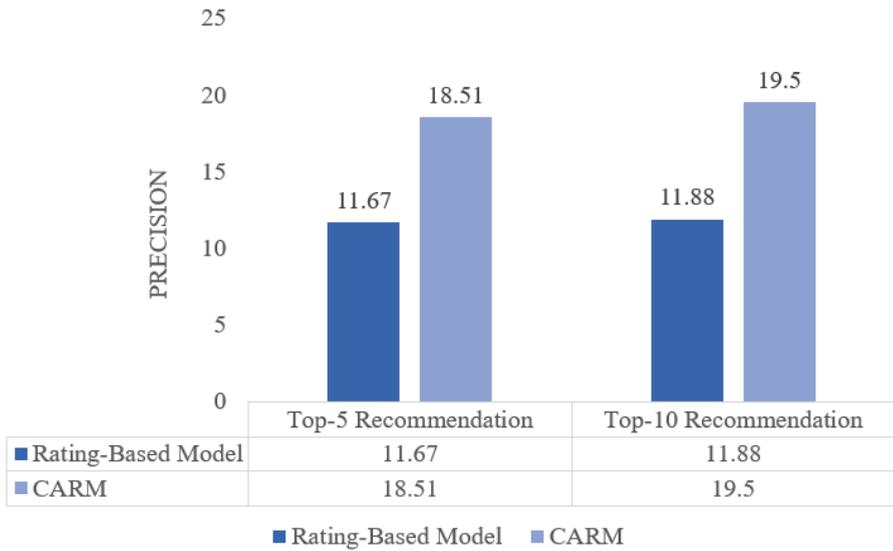


Fig. 2 Performance Evaluation of CARM using Precision

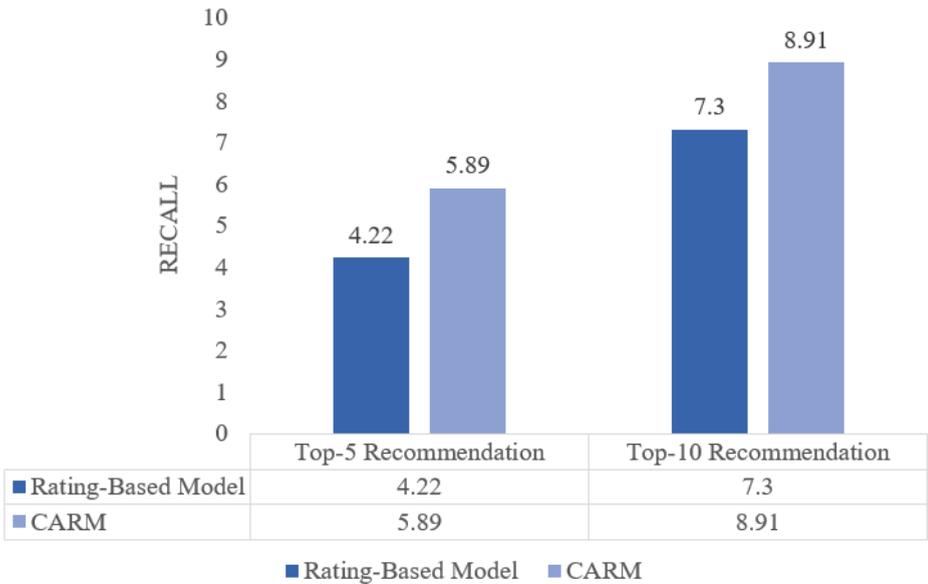


Fig. 3 Performance Evaluation of CARM using Recall

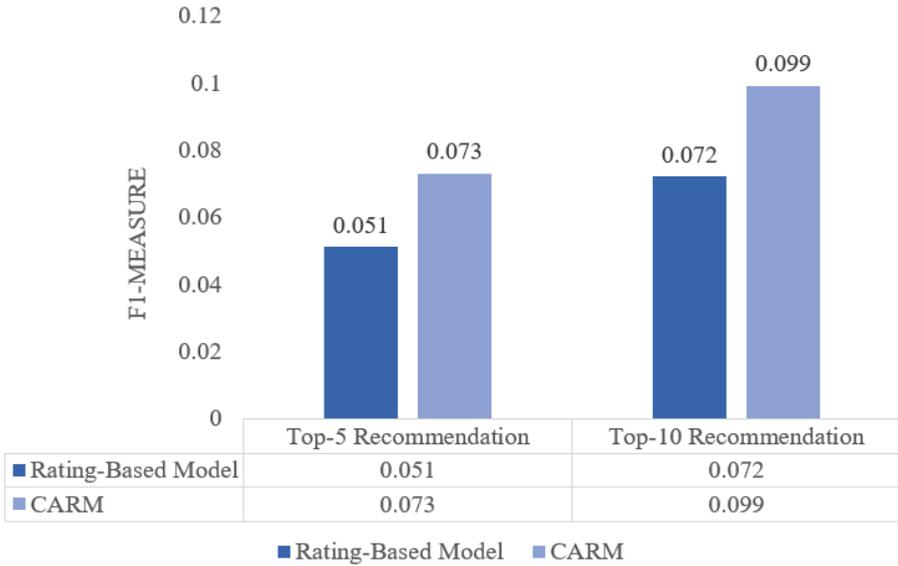


Fig. 4 Performance Evaluation of CARM using F1-Measure

Table 5. Performance Evaluation of the traditional model of CARM on Yelp Dataset using MAE and RMSE

Methods	MAE	RMSE
Rating-Based Model	0.6904	0.8950
Proposed Model	0.6402	0.8612

8 Conclusion

With the increase in online information available to users, innovative methods are needed to assist the user in finding relevant information. Our paper shows how a hybrid approach and sentimental analysis can be put within a common framework to have more precise recommendations. Here, we propose a context-aware restaurant recommender system to incorporate context information and user preferences in the restaurant recommender system. The model first extracts the user preferences by preprocessing, clustering, and sentiment analysis. Later, we find the relevant contextual information in our model and incorporate it with user preference to the recommendation engine. Lastly, the top-n list of restaurants is recommended. The recommendation is based on their similarity of user preferences and contextual information obtained through the proposed model.

References

1. Z. D. Champiri, S. R. Shahamiri and S. S. B. Salim, "A systematic review of scholar context-aware recommender systems," *Expert Systems with Applications*, vol. 42, no. 3, pp. 1743-1758, 2015.
2. A. K. Dey, "Understanding and using context," *Personal and ubiquitous computing*, vol. 5, no. No. 1, pp. 4-7, 2001.
3. Bolchini, Cristiana, F. A. Schreiber and L. Tanca, "A methodology for a very small database design," *Information Systems*, vol. 32, no. No. 1, pp. 61-82, 2007.
4. J.-y. Hong, E.-h. Suh and S.-J. Kim, "Context-aware systems: A literature review and classification," *Expert Systems with applications*, vol. 36, no. No. 4, pp. 8509-8522, 2009.
5. B. N. Schilit and M. M. Theimer, "Disseminating active map information to mobile hosts," *IEEE network*, vol. 8, no. 5, pp. 22-32, 1994.
6. Brown, J. Peter, J. D. Bovey and X. Chen, "Context-aware applications: from the laboratory to the marketplace," *IEEE personal communications*, vol. 4, no. No. 5, pp. 58-64, 1997.
7. Abowd, G. D., A. K. Dey, P. J. Brown, N. Davies, M. Smith and P. Steggles, "Towards a better understanding of context and context-awareness," In *International symposium on handheld and ubiquitous computing*, Vols. Springer, Berlin, Heidelberg, pp. 304-307, 1999.
8. P. Dourish, "What we talk about when we talk about context," *Personal and Ubiquitous Computing*, vol. 8, no. No.1, pp. 19-30, 2004.
9. B. N. Schilit, N. Adams and R. Want, "Context-aware computing applications," In *1994 first workshop on mobile computing systems and applications*, vol. IEEE, pp. 85-90, 1994.
10. Haseloff and Sandra, "Context awareness in information logistics," 2005.
11. Z. Abbasi-Moud, H. Vahdat-Nejad and J. Sadri, "Tourism recommendation system based on semantic clustering and sentiment analysis," *Expert Systems with Applications*, vol. 167, no. 114324, 2021.
12. A. K. Dey and G. D. Abowd, "Towards a better understanding of context and context-awareness," In *International symposium on handheld and ubiquitous computing*, Vols. Springer, Berlin, Heidelberg, pp. 304-307, 1999.
13. L. Baltrunas, "Exploiting contextual information in recommender systems," in *In Proceedings of the 2008 ACM conference on Recommender systems* (pp. 295-298), 2008.
14. G. Adomavicius, R. Sankaranarayanan, S. Sen and A. Tuzhilin, "Incorporating contextual information in recommender systems using a multidimensional approach," *ACM Transactions on Information systems (TOIS)*, vol. 23, no. no. 1, pp. 103-145, 2005.
15. B. Liu and L. Zhang, "A survey of opinion mining and sentiment analysis," *Mining text data*. Springer, Boston, MA., pp. 415-463, 2012.
16. D. A. Pereira, "A survey of sentiment analysis in the Portuguese language," *Artificial Intelligence Review*, vol. 54, no. 2, pp. 1087-1115, 2021.
17. D. M. E.-D. M. Hussein, "A survey on sentiment analysis challenges," *Journal of King Saud University-Engineering Sciences*, vol. 30, no. 4, pp. 330-338, 2018.
18. R. R. Subramanian, N. Akshith, G. N. Murthy, M. Vikas, S. Amara and K. Balaji, "A Survey on Sentiment Analysis," *1th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pp. 70-75, 2021.

19. C. Anderson, "A survey of food recommenders," arXiv preprint arXiv:1809.02862, 2018.
20. S. B. Hegde, S. Satyappanavar and S. Setty, "Sentiment based food classification for restaurant business," 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), no. IEEE, pp. 1455-1462, 2018.
21. M. Ashok, S. Rajanna, P. V. Joshi and S. k. S, "A Personalized Recommender System using Machine," 2016 IEEE Students' Conference on Electrical, Electronics and Computer Science (SCEECS), 2016.
22. C. Musto, M. d. Gemmis, G. Semeraro and P. Lops, "A multi-criteria recommender system exploiting aspect-based sentiment analysis of users' reviews," in Proceedings of the eleventh ACM conference on recommender systems, 2017.
23. M. Deac-Petrusel and S. Limboi, "A sentiment-based similarity model for recommendation systems," 2020 22nd International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC), vol. IEEE, pp. 224-230, 2020.
24. K. Jalan and P. K. Gawande, "Context-Aware Hotel Recommendation System based on Hybrid Approach to Mitigate Cold-Start-Problem," International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS-2017), no. IEEE 2017, pp. 2364-2370, 2017.
25. T. T. A. Nguyen and S. Conrad, "A Semantic Similarity Measure between Nouns based on the Structure of Wordnet," In Proceedings of International Conference on Information Integration and Web-based Applications & Services, pp. 605-609, 2013.
26. B. Vargas-Govea, G. González-Serna and R. Ponce-Medellín, "Effects of relevant contextual features in the performance," ACM RecSys 11, vol. no. 592, p. 56, 2011.
27. G. Capelleveen, C. Amrit, D. M. Yazan and H. Zijm, "The recommender canvas: a model for developing and documenting recommender system design.," Expert systems with applications, vol. 129, pp. 97-117, 2019.
28. S. Baccianella, A. Esuli and F. Sebastiani, "Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining," In Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10), 2010.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

