Forecasting movie rating using k-nearest neighbor based collaborative filtering

Prakash Pandharinath Rokade, PVRD Prasad Rao, Aruna Kumari Devarakonda
[1]Department of Computer Science and Engineering, KLEF Deemed University, Vaddeswaram, India

Article Info

Article history:

Received May 13, 2021
Revised Jun 6, 2022
Accepted Jul 1, 2022

Keywords:

Collaborative filtering
K-nearest neighbor
Machine learning
Recommended system

ABSTRACT

Expressing reviews in the form of sentiments or ratings for item used or movie seen is the part of human habit. These reviews are easily available on different social websites. Based on interest pattern of a user, it is important to recommend him the items. Recommendation system is playing a vital role in everyone's life as demand of recommendation for user's interest increasing day by day. Movie recommendation system based on available ratings for a movie has become interesting part for new users. Till today, a lot many recommendation systems are designed using several machine learning algorithms. Still, sparsity problems, cold start problem, scalability, grey sheep problem are the hurdles for the recommendation systems that must be resolved using hybrid algorithms. We proposed in this paper, a movie rating system using a k-nearest neighbor (KNN-based) collaborative filtering (CF) approach. We compared user's ratings for different movies to get top K users. Then we have used this top K set to find missing ratings by user for a movie using CF. Our proposed system when evaluated for various criteria shows promising results for movie recommendations compared with existing systems.

Corresponding Author:

Prakash Pandharinath Rokade
Department of Computer Science and Engineering, KLEF Deemed University
Vaddeswaram, India
Email: prakashrokade2005@gmail.com

1. INTRODUCTION

Now a days internet has become an integral part of life. People are strongly connected to social media for sharing their emotions and reviews on different websites. These emotions are in the form of sentiments or ratings for a product or service. Business intelligence for provider or demander is guided by recommendation system. Recommendation system suggests items or services to the user of his interest like books, movies, shares. A large dataset with ratings by different users is available online. The movie rating database with users and different parameters for movies is available on several popular websites like Kaggle. Decisions made by support of multiple stronger historical impressions to resolve an issue are always superior to the decisions made with single impression by any user. Rather than collecting all of the reviews or ratings, only the users having stronger relevance of ratings between them are collected. By using k-nearest neighbor (KNN) recommendation system finds a small set of data which is the most similar.

Collaborative filtering (CF) takes input from KNN clustering algorithm. KNN gives top K users with ratings among available huge dataset. Movies genres may be of entertainment, educational, horror, and comedy. Movie recommendation system will help user to rate unrated or new movie for which old users have given ratings already. The increasing amount of data or internet raised challenges for the users to manage the information for getting suggestion to them to predict the rating for their interest.

2. LITERATURE REVIEW

Godbole *et al.* [1] proposed a mechanism for assigning positive or negative scores to each entity in a text collection. Annett and Kondrak [2] discuss an innovative application of support vector machines (SVM). They used movie data to compare their method to lexical-based [2]. The results of different classifiers suggest that using numerous classifiers in a hybrid approach increases sentiment classification quality [3]. For sentiment analysis (SA) of reviews for travel sentiments, three alternative machine learning methods are compared [4]. SentiWordNet is proposed as a source for building data set, and advancement is proposed by calculating positive and negative scores of phrases to identify sentiment orientation [5].

Shambour and Lu [6] addressed how e-governments assist in the search for a reliable corporate partner to conduct business. Moshfeghi *et al.* [7] discuss a novel strategy for CF that uses a combined approach of semantic, emotional, and rating information. The product features for which an opinion is offered are mined in an opinion summary, which varies from text summarization [8]. By clients Erik Cambria outlined how SA may be used in recommendation systems in great detail. The comment is replaced by a rating, followed by a discussion of rating approach [9]. The influence of domain information is considered when choosing feature vectors. Different classifiers are used to determine their impact on a specific domain and feature vector [10]. Pappas and Popescu-Belis [11] proposed a sentiment-aware closest neighbor algorithm for multimedia recommendations based on user comments. A high-prediction-accuracy explicit factor model for explainable suggestions is proposed [12]. Tang *et al.* [13] discusses how feature-based twitter sentiment is superior than standard neural networks. A novel memory-based CF technique is suggested, which models a CF-based recommender system using user reviews and numeric ratings [14].

Yang *et al.* [15] proposed a clique-based data smoothing approach to solve the data sparsity problem in CF using traditional user based nearest neighbor algorithm. Trends and patterns in client priorities have been studied by Subramaniyaswamy *et al.* [16] can be used with filtering and clustering techniques to identify items of interest. Nilashi *et al.* [17] presents a dimensionality reduction strategy to overcome the sparsity and scalability problems in the recommended system. The merits and limitations of basic sentiment analysis concepts are thoroughly explained and compared [18]. On several datasets, Özdemir *et al.* [19] demonstrated how to evaluate the efficacy of various data classifier algorithms. Mohemad *et al.* [20] proposed the construction of a new ontology model in the education sector in order to give early detection of children with learning problems.

Babeetha *et al.* [21] proposed a prediction strategy for smoothing sparse original rating matrices and clustering, as well as a discussion of the proposed methods' accuracy and processing time. Rahim *et al.* [22] looked into the importance of a number of important variables for innovative digital marketing. Heart disease and breast cancer can be predicted. Saranya and Pravin [23] worked on developing successful prediction models and methodologies. Behera *et al.* [24] described how collaborative filtering and content-based recommendation are well-known strategies for selecting movies from a huge collection and determining their attribute-based similarity. According to Ez-Zahout *et al*. [25] matrix factorization is the superior technique, and it delivers a satisfactory result with a high precision score for the MovieLens dataset. Mawane *et al*. [26] suggested a platform whose purpose is to try to discover the suitable parameters of the Kohonen maps that can help satisfy the relevance of recommended items by dividing learners into homogeneous groups before generating recommendations. El Fazziki *et al*. [27] examined user-based CF on two datasets, film trust and MovieLens, and found that it works well and enhances prediction accuracy. Rawat *et al.* [28] proposed a method for predicting virtual machine failure based on a time series stochastic model that accurately predicts failure. Various evaluation techniques like precision, recall, F-measure, accuracy for machine learning based algorithm are discussed by Yadla and Rao [29].

3. RESEARCH CHALLENGES AND OBJECTIVES

Today, several recommender systems have been evolved for specific domains but, those are not particular enough to fulfill the information desires of users. So, it is important to build better recommender system. In designing such system, we face numerous issues.

3.1. Sparsity problems

A recommendation quality depends on data sparsity. Data sparsity occurs when dimensionality of the data set increases and some users do not rate few items. Good research is carried out to minimize this problem. Dimensionality reduction can solve this issue up to a certain extent.

3.2. Cold start problem

The cold start problem occurs as new users or items enter the system. This issue is classified as a new user issue, a new item issue, or a new system issue. This issue degrades the performance of collaborative filtering for recommendation. We may look into appropriate user group rather than entire data set-to resolve this issue.

3.3. Scalability

Small dataset can be evaluated to recommend users for their choices. As dataset increases, accuracy in result reduces. We can use advanced large scale assessment methods to improve scalability.

3.4. Grey sheep problem

Sometimes user's behavior is suspicious as the user may rate related items inconsistently. This will degrade the performance of recommendation system. One way omits these users from training dataset and choose only top K users we have rated all items with consistency to related items. Objectives of work: i) to analyze the given ratings by different users for movies; ii) to minimize sparsity, cold start issue; and iii) to compare proposed approach with existing approaches.

4. PROPOSED METHOD ARCHITECTURE

As shown in Figure 1, the input dataset is of Movielens from Kaggle database. KNN algorithm is applied on this dataset to find Top K similar users who have rated all movies. Collaborative filtering algorithm is applied on this Top K dataset for recommending rating for unrated movie.

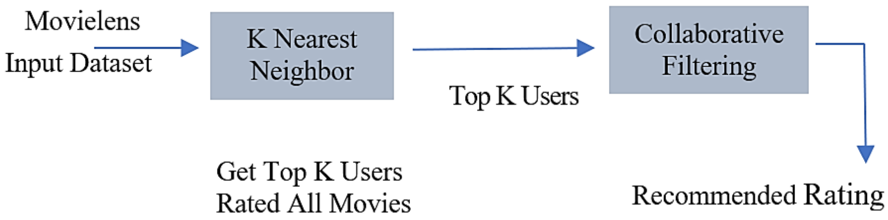

Figure 1. The proposed system's architecture

4.1. Input dataset

The input dataset contains following features: i) user Id; ii) movie Id; and iii) rating. In the Table 1, the record set contains ratings given by various users to various movies. Six users have rated 4 movies. The ratings given by users vary from 1 to 5. The record set additionally contains few rows where rating data is missing, which will be computed using CF. To find missing rating, we first find top K users having higher similarity between them.

Table 1. Ratings for movies by users

User Vs Movie	1	2	3	4
A	3	4.5	2	4
B	4	1	2	5
C	3	4.5	2	4
D	5	4.5	3	1
E	3	1	5	5
X	3	2	3	

4.2. Finding top K users

Here our aim is to find top K similar users whose ratings for all movies are closer than remaining users. To attain this, we are going to use KNN clustering algorithm. K in KNN stands for the number of neighbors. In training period KNN does not learn anything. It saves the training data set and uses it to make real-time predictions. Due to this KNN algorithm is much faster than support vector machine, regression which require training before prediction. To implement KNN we require value of K and distance function. We are going to use Euclidean distance function to find top K users amongst all used set. KNN algorithm: i) select the K users who have given ratings to all movies, ii) find the Euclidean distance of each user with remaining all users, iii) select top K users with maximum weight to formulate a set of top K users, and iv) our model is ready.

4.3. Formula for Euclidean distance

To calculate the distance between user A and B, we will put values of their ratings in (1). Ratings of user A will occupy positions of variable x whereas ratings of user B will occupy positions of variable y. For calculating the Euclidian distances, we will consider all the users who have rated all movies. As shown in Table 2, first we will find the average rating by every user.

Table 2. Average ratings by users

User Vs Movie	1	2	3	4	Average Rating
A	3	4.5	2	4	3.16
B	4	1	2	5	2.33
C	3	4.5	2	4	3.16
E	3	1	5	5	3

$$D(x,y) = \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + \ldots .(xn - yn)^2} \quad (1)$$

By putting the values from Table 2 in the formula we will get the Euclidean distance between *A* and *B*.

$$D(A,B) = \sqrt{(3 - 4)^2 + (4.5 - 1)^2 + (3 - 2)^2 + (4 - 5)^2}$$

$$D(A,B) = 3.9$$

Similarly, by calculating Euclidean distances for user pairs (A,C), (A,D), (A,E), (B,C), (B,D), (B,E), (C,D), (C,E), (D,E) we will get following results shown in Table 3. As shown in Table 4, if we consider K=1 we will get smallest 1 distance that is 0 for (A,C), for K=2 we will get smallest 2 distances 0 and 3.16 for (A,C) and (B,E) respectively. Similarly for K=3 we will get smallest 3 distances 0 for (A,C), 3.16 for (B,E), 3.74 for (A,D) and (C,D).

Table 3. Euclidean distance between users

User pair	Distance between users
A,B	3.9
A,C	0
A,D	3.74
A,E	4.71
B,C	3.77
B,D	5.5
B,E	3.16
C,D	3.74
C,E	4.71
D,E	6.02

Table 4. User pair set for different K values

K value	Distance	User pairs
K=1	0	(A,C)
K=2	0,3.16	(A,C),(B,E)
K=3	0,3.16,3.74	(A,C),(B,E),(A,D),(C,D)

If we consider value of K very small, we will face the problem of under fitting in which noisy data or outlier has huge impact on our classifier. In opposite case, if we consider value of K very large, we will face the problem of over fitting in which classifier has tendency to predict majority classes regardless of which neighbors are nearest. To avoid under fitting and over fitting we will consider value of K=2. Then we will get following in Table 5. In Table 5 only the users by considering value of K=2 is shown with their ratings.

Table 5. User with ratings for K=2

K value	Distance	User Pairs
K=1	0	(A,C)
K=2	0,3.16	(A,C), (B,E)

4.4. Collaborative filtering

The aim of CF is to predict ratings for unrated items by considering the top K user set with their ratings. We will not consider the column with missing ratings. As shown in Table 5, we will ignore the column for movie 4 as user X has not rated it. Now we will find average of all ratings provided by all users for movie 1, 2, 3. The average ratings are shown in Table 6. Now we will find the similarity between user X with all remaining users. We will attain this by cosine similarity formula.

Table 6. Average ratings by users without considering missing value column

User Vs Movie	1	2	3	4	Average Rating
A	3	4.5	2	4	3.16
B	4	1	2	5	2.33
C	3	4.5	2	4	3.16
E	3	1	5	5	3
X	3	2	3		2.66

4.5. Comparing the similarity between users

Now, we will compare ratings of user A, B, C, E with the ratings of user X to get how close these users are with user X. To achieve this, we preferred to use cosine similarity formula. Cosine similarity is a metric, helpful in determining, how similar the data objects are irrespective of their size. Cosine similarity formula:

$$\mathrm{sim}(a,b) = \frac{\vec{a}.\vec{b}}{|\vec{a}|.|\vec{b}|} \tag{2}$$

$$\mathrm{sim}(C_i, C_j) = \frac{\Sigma[(r_{ip}-r_{i\,avg})-(r_{jp}-r_{j\,avg})]}{\sqrt{\Sigma(r_{ip}-r_{i\,avg})^2}\sqrt{\Sigma(r_{jp}-r_{j\,avg})^2}} \tag{3}$$

where, r_{ip} is current/particular rating of customer i, r_{jp} is current/particular rating of customer j, r_{iavg} is average rating of customer I and r_{javg} is average rating of customer j.

The similarity varies in between -1 to 1 only. To calculate similarity between user A and X we will put values in (3).

$$\mathrm{sim}(A,X) = \frac{\begin{array}{c}(3-3.16)(3-2.66)+(4.5-3.16)(2-2.66)\\+(2-3.16)(3-2.66)\end{array}}{\begin{array}{c}(\sqrt{(3-3.16)^2+(4.5-3.16)^2+(2-3.16)^2})\\ * \\ (\sqrt{(3-2.66)^2+(2-2.66)^2+(3-2.66)^2})\end{array}}$$

$$= \frac{(-0.16)(0.34)+(1.34)(-0.66)+(-1.66)(1.66)}{\sqrt{(-0.16)^2+(1.34)^2+(-1.16)^2}\ \sqrt{(0.34)^2+(-0.66)^2+(1.66)^2}} = -0.89$$

Similarly, by calculating cosine similarity for (B,X), (C,X) and (E,X) we get results as shown in Table 6.

From Table 7, we will get lowest similarity weight -0.89 for user pair (A,X) whereas highest similarity weight 0.31 for user pair (E.X). Hence as similarity between ratings of pair (E,X) is highest; we can assign user E's rating for movie 4 to missing rating for movie 4 by user X. So, the missing rating for user X for movie 4 will be 5.

Table 7. Cosine similarity between user pairs

Sim (Ci,Cj)	Similarity weight	Remark
(A,X)	-0.89	Lowest similarity
(B,X)	0.22	Moderate similarity
(C,X)	-0.28	Moderate similarity
(E,X)	0.31	**Highest similarity**

5. CONCLUSION

In this paper, we proposed a recommended system based on collaborative filtering with k-nearest neighbors (KNN). We got unknown rating by comparing the ratings for top k-movies. Selection of proper value of K improved our result of prediction. We have resolved both the issues of underfitting and overfitting by selecting proper K value. In future, we should use deep learning techniques for movie rating prediction or movie recommendations. Also, we may prefer some standard database directly from the organizations working in the same domain to get more accurate results.

REFERENCES

[1] N. Godbole, M. Srinivasaiah, and S. Skiena, "Large scale sentiment analysis for news and blogs," *ICWSM'2007 Boulder*, 2007.

[2] M. Annett and G. Kondrak, "A comparison of sentiment analysis techniques: polarizing movie blogs," in *Advances in Artificial Intelligence*, Springer Berlin Heidelberg, 2008, pp. 25–35, doi: 10.1007/978-3-540-68825-9_3.

[3] R. Prabowo and M. Thelwall, "Sentiment analysis: A combined approach," *Journal of Informetrics*, vol. 3, no. 2, pp. 143–157, Apr. 2009, doi: 10.1016/j.joi.2009.01.003.

[4] Q. Ye, Z. Zhang, and R. Law, "Sentiment classification of online reviews to travel destinations by supervised machine learning approaches," *Expert Systems with Applications*, vol. 36, no. 3, pp. 6527–6535, Apr. 2009, doi: 10.1016/j.eswa.2008.07.035.

[5] B. Ohana and B. Tierney, "Sentiment classification of reviews using SentiWordNet," *9th. IT &T Conference*, 2009.

[6] Q. Shambour and J. Lu, "A hybrid trust-enhanced collaborative filtering recommendation approach for personalized government-to-business e-services," *International Journal of Intelligent Systems*, vol. 26, no. 9, pp. 814–843, Sep. 2011, doi: 10.1002/int.20495.

[7] Y. Moshfeghi, B. Piwowarski, and J. M. Jose, "Handling data sparsity in collaborative filtering using emotion and semantic based features," *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information-SIGIR '11*, 2011, doi: 10.1145/2009916.2010001.

[8] G. Vinodhini and R. Chandrasekaran, "Sentiment analysis and opinion mining: a survey," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 2, no. 6, pp. 282–292, 2012.

[9] M. Á. García-Cumbreras, A. Montejo-Ráez, and M. C. Díaz-Galiano, "Pessimists and optimists: Improving collaborative filtering through sentiment analysis," *Expert Systems with Applications*, vol. 40, no. 17, pp. 6758–6765, Dec. 2013, doi: 10.1016/j.eswa.2013.06.049.

[10] IEEE, "IEEE standard for wireless LAN medium access control (MAC) and physical layer (PHY) specifications," *IEEE Std 802.11-1997*. IEEE, pp. 1–445, 1997, doi: 10.1109/ieeestd.1997.85951.

[11] N. Pappas and A. Popescu-Belis, "Sentiment analysis of user comments for one-class collaborative filtering over ted talks," in *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval*, Jul. 2013, pp. 773–776, doi: 10.1145/2484028.2484116.

[12] Y. Zhang, G. Lai, M. Zhang, Y. Zhang, Y. Liu, and S. Ma, "Explicit factor models for explainable recommendation based on phrase-level sentiment analysis," in *Proceedings of the 37th international ACM SIGIR conference on Research and development in information retrieval*, Jul. 2014, pp. 83–92, doi: 10.1145/2600428.2609579.

[13] D. Tang, F. Wei, N. Yang, M. Zhou, T. Liu, and B. Qin, "Learning sentiment-specific word embedding for twitter sentiment classification," in *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1:Long Papers)*, 2014, pp. 1555–1565, doi: 10.3115/v1/P14-1146.

[14] Z. Zhang, D. Zhang, and J. Lai, "urCF: user review enhanced collaborative filtering," *Twentieth Americas Conference on Information System, Savannah*, 2014

[15] Y. Yang, Z. Zhang, and X. Duan, "Cliques-based data smoothing approach for solving data sparsity in collaborative filtering," *TELKOMNIKA Indonesian Journal of Electrical Engineering*, vol. 12, no. 8, Aug. 2014, doi: 10.11591/telkomnika.v12i8.4617.

[16] V. Subramaniyaswamy, R. Logesh, M. Chandrashekhar, A. Challa, and V. Vijayakumar, "A personalised movie recommendation system based on collaborative filtering," *International Journal of High Performance Computing and Networking*, vol. 10, no. 1/2, 2017, doi: 10.1504/IJHPCN.2017.083199.

[17] M. Nilashi, O. Ibrahim, and K. Bagherifard, "A recommender system based on collaborative filtering using ontology and dimensionality reduction techniques," *Expert Systems with Applications*, vol. 92, pp. 507–520, Feb. 2018, doi: 10.1016/j.eswa.2017.09.058.

[18] P. P. Rokade and A. K. D, "Business intelligence analytics using sentiment analysis-a survey," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 1, pp. 613–620, Feb. 2019, doi: 10.11591/ijece.v9i1.pp613-620.

[19] A. Özdemir, U. Yavuz, and F. A. Dael, "Performance evaluation of different classification techniques using different datasets," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 5, pp. 3584–3590, Oct. 2019, doi: 10.11591/ijece.v9i5.pp3584-3590.

[20] R. Mohemad, N. F. A. Mamat, N. M. M. Noor, and A. C. Alhadi, "The development of an ontology model for early identification of children with specific learning disabilities," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 6, pp. 5486–5494, Dec. 2019, doi: 10.11591/ijece.v9i6.pp5486-5494.

[21] S. Babeetha, B. Muruganantham, S. G. Kumar, and A. Murugan, "An enhanced kernel weighted collaborative recommended system to alleviate sparsity," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 1, pp. 447–454, Feb. 2020, doi: 10.11591/ijece.v10i1.pp447-454.

[22] H. A. Rahim, S. Ibrahim, S. B. A. Kamaruddin, N. A. Md. Ghani, and I. Musirin, "Exploration on digital marketing as business strategy model among Malaysian entrepreneurs via neurocomputing," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 9, no. 1, pp. 18–24, Mar. 2020, doi: 10.11591/ijai.v9.i1.pp18-24.

[23] G. Saranya and A. Pravin, "A comprehensive study on disease risk predictions in machine learning," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 4, pp. 4217–4225, Aug. 2020, doi: 10.11591/ijece.v10i4.pp4217-4225.

[24] D. K. Behera, M. Das, S. Swetanisha, and P. K. Sethy, "Hybrid model for movie recommendation system using content K-nearest neighbors and restricted Boltzmann machine," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 23, no. 1, pp. 445–452, Jul. 2021, doi: 10.11591/ijeecs.v23.i1.pp445-452.

[25] A. Ez-zahout, H. Gueddah, A. Nasry, R. Madani, and F. Omary, "A hybrid big data movies recommendation model based k-nearest neighbors and matrix factorization," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 26, no. 1, pp. 434–441, Apr. 2022, doi: 10.11591/ijeecs.v26.i1.pp434-441.

[26] J. Mawane, A. Naji, and M. Ramdani, "A cluster validity for optimal configuration of Kohonen maps in e-learning recommendation," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 26, no. 1, pp. 482–492, Apr. 2022, doi: 10.11591/ijeecs.v26.i1.pp482-492.

[27] A. El Fazziki, Y. E. M. El Alami, J. Elhassouni, O. El Aissaoui, and M. Benbrahim, "Employing opposite ratings users in a new

approach to collaborative filtering," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 25, no. 1, pp. 450–459, Jan. 2022, doi: 10.11591/ijeecs.v25.i1.pp450-459.
[28] A. Rawat, R. Sushil, A. Agarwal, and A. Sikander, "A new approach for VM failure prediction using stochastic model in cloud," *IETE Journal of Research*, vol. 67, no. 2, pp. 165–172, Mar. 2021, doi: 10.1080/03772063.2018.1537814.
[29] H. K. Yadla and P. P. Rao, "Machine learning based text classifier centered on TF-IDF vectoriser," *Journal of Scientific & Technology Research*, vol. 9, no. 3, pp. 583–586, 2020.

BIOGRAPHIES OF AUTHORS

Prakash Pandharinath Rokade graduated from Pune University in Maharashtra, India, with a bachelor's degree in computer engineering in 2005. In 2011, he obtained his M.Tech. in Computer Engineering at Bharti Vidyapeerth Pune, Maharashtra, India. He is currently pursuing his Ph.D. in Computer Science and Engineering at Koneru Lakshmaiah Education Foundation, previously K L University in Vaddeswaram, Andhra Pradesh, India. Sentiment analysis, opinion mining, and machine learning are some of his research interests. He can be contacted at email: prakashrokade2005@gmail.com.

PVRD Prasad Rao has 23 years of experience in industry and academia. In the field of Data Science, he obtained his M.Tech in CSE from Andhra University in 1998 and his Ph.D. in Computer Science from Acharya Nagarjuna University in 2014. He is currently employed as a Professor at KL University's Department of Computer Science and Engineering. Data science, bioinformatics, IoT, and cyber security are among his research interests. He has around 70 research publications published in journals and conference proceedings from SCI, WoS, and Scopus. He is also an Associate Dean (P&P), a Reviewer, and an Editorial member for a number of journals. He can be contacted at email: pvrdprasad@kluniversity.in.

Aruna Kumari Devarakonda earned her Ph.D. in Computer Science and Engineering from K L University in Vaddeswaram, Andhra Pradesh, India. Professor at Koneru Lakshmaiah Education Foundation, formerly K L University, is her current position. Her teaching and research interests include data mining and machine learning, and she has over 50 publications published in national and international journals. The DST Young Scientist Award has been given to her (Government of India). She can be contacted at email: arunakumari@vjit.ac.in.