

Research Article

Adaptive and Personalized Web Blog Searching Technique Using S-ANFIS

Harsh Khatter ^{1,2} **Pooja Malik** ³ **Amrita Jyoti** ⁴ **Anil Kumar Ahlawat** ^{2,5}
Gaurav Dubey ^{1,2} **Anurag Mishra** ^{1,2} and **Sanjeev Chandra Neupane** ⁶

¹Department of Computer Science, KIET Group of Institutions, Delhi-NCR, Ghaziabad, India

²Dr. A.P.J. Abdul Kalam Technical University, Lucknow, India

³Department of Computer Science and Engineering, School of Engineering, Shiv Nadar University, India

⁴Department of Information Technology, Ajay Kumar Garg Engineering College, Ghaziabad, India

⁵KIET Group of Institutions, Delhi-NCR, Ghaziabad, India

⁶Data Strategist, AV Digitals, Pulchowk, Lalitpur, Nepal

Correspondence should be addressed to Sanjeev Chandra Neupane; sanjeev@reconwithme.com

Received 17 December 2021; Accepted 10 August 2022; Published 27 August 2022

Academic Editor: Pradeep Kumar Singh

Copyright © 2022 Harsh Khatter et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Day by day, the number of blog users and microblog users is increasing worldwide. It is easy to say that blogs have captured a significant portion of other web services. In the past few years, the number of users has exponentially increased. User count of Facebook, Twitter, and Instagram applications is not hidden from anyone. Users on such platforms share ideas, experiences, stories, opinions, and views and want to interact with people with the same set of interests. As per the user's expectation, there is a requirement of two things: *content curation* and *recommendations*. The content curation algorithm will find the people and their posts on personalized search results. In addition, the recommendation system will help to find the most appropriate match to interact with. In this paper, both approaches are combined to show the user's curated and recommended results. The article focuses on the hybrid model named S-ANFIS, and the results are compared with the well-known approaches like ANN, Deep Neural Network (DNN), and Recurrent Neural Network (RNN).

1. Introduction

Under the umbrella of Web 2.0, a blog is one of the essential tools nowadays. People are writing their blogs to share their ideas, thoughts, creativity, and articles with others. Blog writers feel privileged to show their point of view on recent trends and news. Many blogs are specially meant for politics, where the blogger shares his idea about government policies and discussions. On the same ground, many economists have their blogs. Based on the user's interests, users create their blogs and enormously use web services to share information worldwide. In the giant world of information, blogs have created their own world called the blogosphere. The more the people share their ideas, the more the creativity spreads. Bloggers have their own interest and insight to view the things and help in

the development of solutions in various domains, i.e., the field of education, technology, science, literature, and many more. Basically, this is the core concept of micro-blogging [1]. Instagram, Twitter, and Facebook are well-known examples of multimedia social networking blogs. Su et al. discussed how the users of such applications want to interact with other users from their locality and similar user preferences [2]. Every day, millions of users are using the services of these platforms and share the posts. On average, each user is associated or followed by hundreds of other users or friends. Practically, is it possible for a user to view the posts shared by all hundreds of connected users? The majority of users say, No. The time which the people are spending on social blogs cannot consider all the updates from all the users. Here, comes the role of personalization [3, 4].

Personalization is one thing that categorizes the posts based on the user's interest, though the friends and followers connected on the platforms are already chosen by the users and their profile. Still, the membership value of each and every instance is calculated and evaluated to show the personalized result to the user. This method helps the user to view the relevant posts in minimum time with minimum scrolling on the screen. In this fast-moving world, researchers are working daily to make the user's life simpler and friendly to the web. The rising popularity of smart phones and such blogging interfaces has given a new rise to collaborative environments [5]. Every day, a new algorithm is coming, which optimizes the use of user's resources and gives the relevant results in a minimum amount of time. In this paper, the content curation algorithm is discussed and applied on the blog posts. The implementation of a fuzzy inference engine with neural network, named S-ANFIS, is shown with the comparative results. The results are taken in three scenarios, i.e., when training data is 70% and testing data is 30%, when training data is 80% and testing data is 20%, and when training data is 90% and testing data is 10%.

2. Literature Survey

Based on the study of related work, the literature work is divided into two parts: first is based on the web services and blogs, and second is based on the working and features of various recommended systems.

2.1. Related Work Based on Web Services and Blogs. Conzanzo and Casas discussed the web frameworks and how web applications and tools help in the evaluation of Web 2.0 and Web 3.0. The Web frameworks are sufficient enough to differentiate among the various platforms and finding out the weaknesses and make improvements in the web tools and services to help the users in finding efficient and accurate results. In the documentation, authors have taken examples of various websites that are related to books, tutorials, and videos, including online communities such as blogs, discussion forums, and online repository like github, stack overflow, wikis, and chats. The analysis of such frameworks has done and determined the usability and quality based on the key attributes, parameters, and measures. They proposed a set of elements for evaluating the usability of support from the analysis of available and published resources of 17 web frameworks, considered among those which are very popular, recognized, and used in web development communities [6].

Jiugen et al. have taken an example of the educational community of China and shared the experience of the thinking of the educational scholars. The authors focused on gratifying the education microblog, which combines student-centered modern education and people-centered thinking. The microblogging defined by Jiugen et al. collaborated educational development with mobile activity learning [1]. The microblogging site, which authors mentioned in their work, is meant for the informal learning process, which helps in information delivery, information exchange, and information sharing. This is the collaborative effort of informal learning resources and Web 2.0 services.

Kaur and Singh highlighted the features of blog and how the blog posts data volume is increasing day by day. The blog sites and the social blog posts are getting popular among the users [7]. Breuch showed the demonstration of the user analysis in two case studies was discussed in the article. The case studies examined the blog web responses from Facebook and Twitter. The comments were explored as social web feedback in relation to website usability. The author has focused on the "audience involvement" to understand the social web feedback about Twitter- and Facebook-type social blog websites [8]. Singh et al. discussed how the social networking services are increasing day by day. The authors did the case study on Facebook web application and analyze the comments and posts. Based on the model shown in Figure 1, the authors tried to predict how many comments a post is expected to receive in the next coming hours. The process contains the web crawler and information extractor in the first phase. Later, the information is processed by information processor module and knowledge discovery module. Major concern was taken care of neural networks and decision trees which performs the predictive modelling technique [9].

Fu et al. studied and discussed the concept of people interaction and detecting groups on microblog sites: how the increasing scope of microblogs is spreading and helps in the development of communication networks among people and how information technology is providing various ways for people to interact with each other. Microblogs are the best way of information dissemination and social interaction to connect the people. People are using the microblogs to express their ideas, thoughts, and feelings. To analyze the reader's summary of blog and the followers of blog, a similar interest group is detected. This analysis helps in understanding the behavior and the evolutionary trend within the community. In their paper, the authors focus on the problem of group detection on a Chinese microblog named Sina. The authors have taken the modified SimHash algorithm to analyze and compute the interest similarity among blog followers. The interest of people followed the blog posts, and the sentiments which the posts expressed in each group were studied by the authors [12]. Ding et al. worked on a similar category of Chinese blog and focused on the semantic analysis of the user's blog. The analysis was done on behalf of the patterns of posts and patterns of tags. The authors applied the random forest algorithm in the study [13]. Wang et al. discussed the ranking of microblogging websites based on a user's tagging. The authors randomly took the sample posts and organized the groups of tags and proposed a user tag ranking schema. The scores were decided based on the relevance among the tags and users to rank the user tags. Experimental results were taken from the dataset of more than 140 million users and gave an excellent performance [14].

2.2. Related Work Based on Recommendation System. Nowadays, we can see the recommendations everywhere while using the web services. It is either an online shopping site, online reservation site, blogging sites, social networking sites, or the content searching sites like Google scholar.

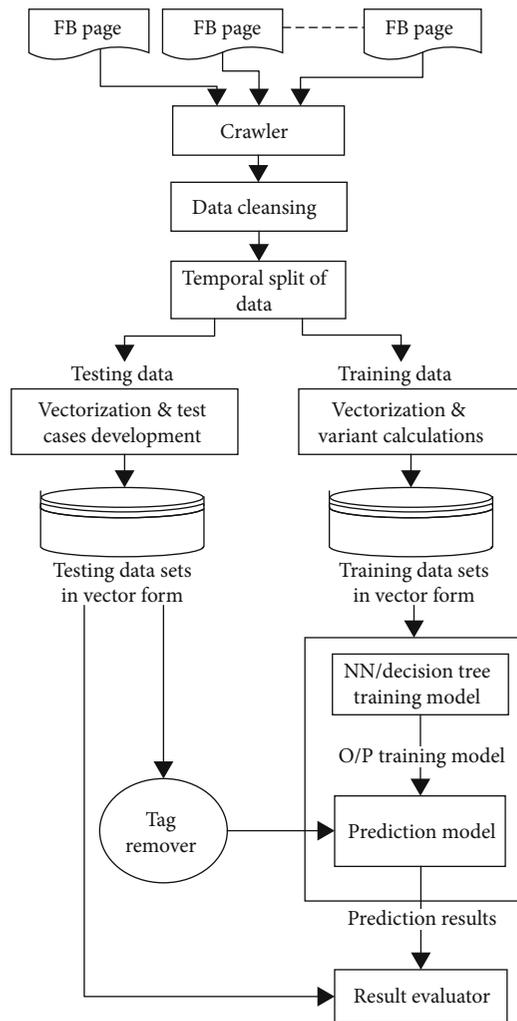


FIGURE 1: Workflow of Facebook [10, 11].

Guo et al. wrote the paper on collaborative filtering and predicting the recommended products to buy to the users based on their profiles. Traditionally, the association rules were applied for the same purpose, now the high-end algorithms are giving the tremendous results, and the predicted values are converted to the actual patterns of buying products [15]. Gao and Wu designed a framework for *Service Set Recommendation in the Mashup Creation*. The authors have taken various approaches and analyze the scope of improving the recommendations. The authors considered two main points: first is the top-ranked services with the same functionality, and second is relations among the services. Based on the list of different aspects, mashup composition patterns were recommended, and experimental evaluation was done on the ProgrammableWeb dataset [16].

Borhanifard and Minaei-Bidgoli gave an idea of recommender system for Persian blogs. The authors collaborate many existing algorithms which fetch and filter the blog posts based on the interests of users and recommend whenever the user logs in. In the paper, a model was discussed which adopts the clustering algorithms and the personalized search results for Persian blogs [17]. Zulkefli and Bin

Baharudin discussed the importance of recommendations in the field of hotels. Based on the user experience, the recommendations are given to the new users. Based on the rating feedback and reviews of the hotel, the recommendations will be displayed to the new users who are seeking to book the hotel [18].

Zhang et al. discussed a recommendation algorithm which is based on dynamic user preference and the service quality. The authors focused on the services provided to the users and how the new services would be recommended to the users based on the past experience. As the time moves on, the user preferences and service quality may change. It means that the need of a user will also change in time. In their research paper, the recommendation algorithm was discussed that considers the dynamic characteristics of users and the dynamic quality of services. The algorithm has taken the base from the temporal LDA (Latent Dirichlet Allocation) model and the quality of service owned by the users. The authors have done the experimental analysis and shown the results based on the real-world dataset [19]. Harrington et al. published a patent about social network-recommended content and the personalized search results. They have explained in a refined manner how the recommendations are generated on web and how the users get benefit from them [10]. Various works are available which use the hybrid fuzzy models on different applications. The desired one is to work on the blog posts and recommend the relevant posts to the user like for recommendations [20–22] highlighted about the user's response on the recommendations by the system, whereas the blog-related work is discussed in [23, 24].

Su et al. discussed a scenario how the friend's suggestions are displayed to the user's screen on social networking sites. The authors have taken an example of friend social networking site and discussed the concept of calculating the matching index based on similar user preferences [2]. Zhou has given an idea for the recommended system for blogs in early days of blog Web 2.0 tool. The author proposed the algorithm based on the personalized user interest and the shared blog information. He discussed the system model and explained the work flow of the retrieved information based on a user's interest. The model is shown in Figure 2 which highlights all the relevant modules of blog recommendation system [25].

Khatteer and Kumar Ahlawat presented an idea to collaborate personalization with the education. This collaboration uses the LMS and comes under e-Learning domain. Similarly, the collaboration of multiple techniques can be considered for better algorithm and results [26].

2.3. Gaps Observed. Based on the literature survey, it is clear that recommendation systems are the backbone of the web services. Recommendation systems help people in exploring the new information based on the user's personalized interest. There are several approaches which can be used to make a recommendation system. Many papers discussed about the clustering algorithm, and some emphasized on the fuzzy logics to make the inference rules to map the user interests. Neural networks and their variants can also be used to design an efficient recommendation system. In this paper,

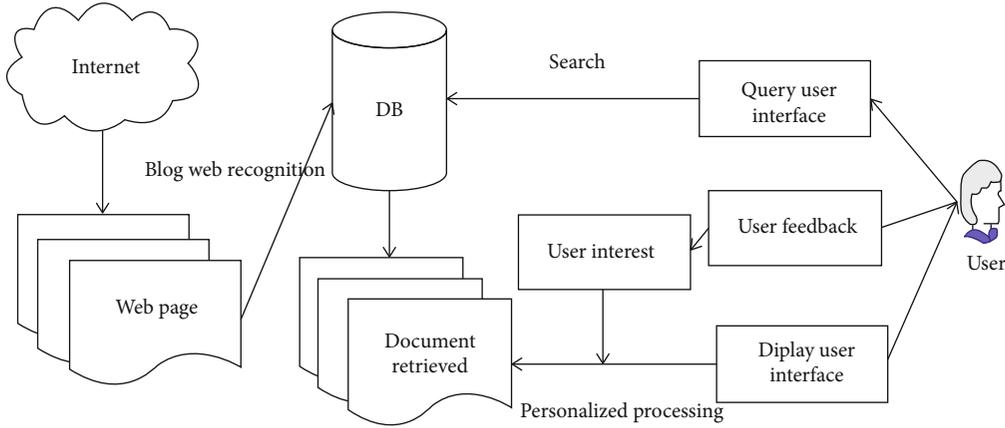


FIGURE 2: Blog web personalized retrieval system architecture [4].

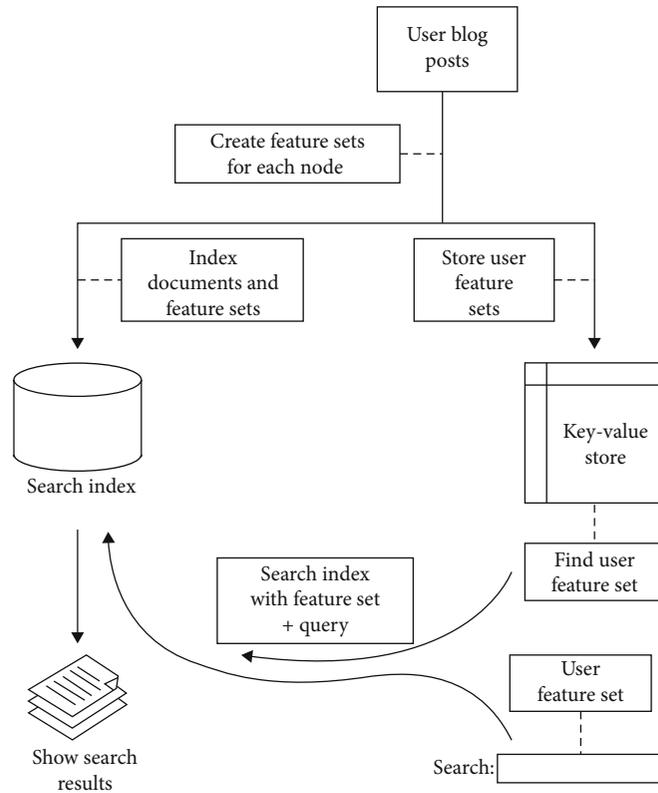


FIGURE 3: Proposed model which collaborates curation and recommendation [9, 11, and 25].

the model and the algorithm for recommendation are discussed which are based on the combination of neural network and fuzzy inference rules, i.e., adaptive neuro-fuzzy inference system (S-ANFIS).

3. Proposed Model

The paper proposed a smart content curation algorithm which comprises four modules. Though the process is defined in many research papers with different names, Rajeswari and Hariharan gave their idea on these four steps in their own words [27].

3.1. Filtering. In the proposed model, the first step is to fetch the data of user in form of blog posts. The key terms are fetched and filtered out to classify the blog users based on their blog post behavior. Some of the key terms are syntactically checked, and some focused on the word semantics.

3.2. Review Analysis. In the previous step, the classes are formed and the division of blog users is done. Review analysis is done on the classes, and the users lie in these classes. Many bloggers post their ideas and views, which are least matched with their own interest. This cross domain information is analyzed in this phase, and a different class is created for this section.

TABLE 1: The outcome of various models implemented on the PIKES dataset.

Model	Training ratio of 0.9			Training ratio of 0.8			Training ratio of 0.7		
	Precision	Accuracy	MAE	Precision	Accuracy	MAE	Precision	Accuracy	MAE
ANN	91.12	92.4	11.4	86	89.4	16.8	84.3	87.4	26.46
DNN	92	93.1	8.2	88	90.2	13.2	86.8	91	18.4
RNN	94.12	95.5	7.6	91	93.4	9.2	91.4	92.8	11.62
S-ANFIS	96.4	96.86	4.2	94.12	95.48	7.8	92.78	93.86	9.4

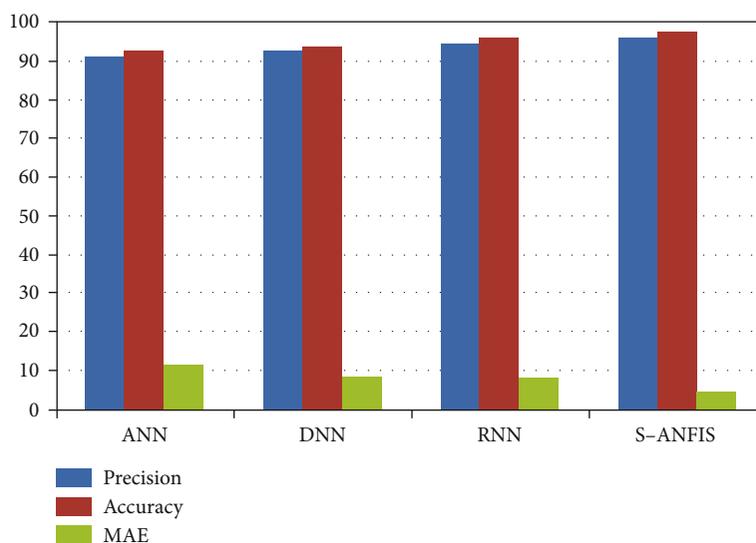


FIGURE 4: Training dataset ratio is 0.9 for PIKES.

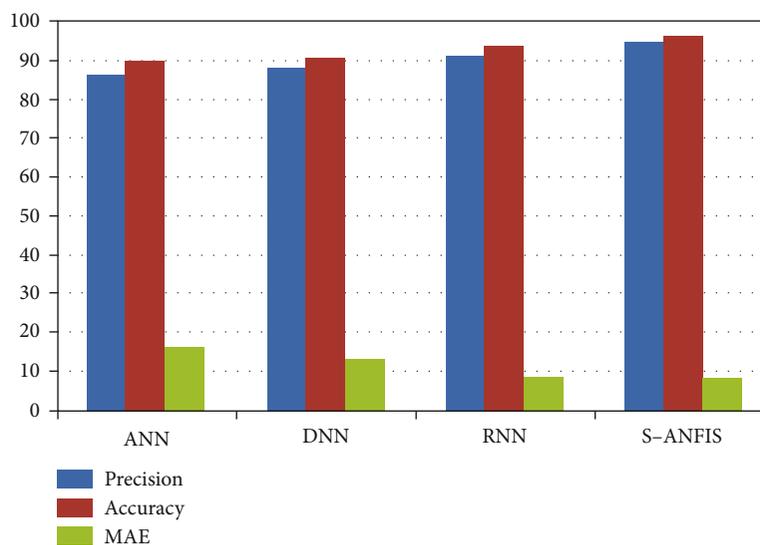


FIGURE 5: Training dataset ratio is 0.8 for PIKES.

3.3. *Similarity.* To create a similarity index among the users of the class, the index values are computed.

3.4. *Recommendation.* The blog users get the recommendation notification only about those bloggers who are carrying the maximum target value computed based on the similarity index S_i . For a particular user i , n number of

the users is filtered out from the same class C . On the similar interest domain, the key terms k were fetched and curated on behalf of user's posts, and the refined terms K are considered. These K terms are applied using the S-ANFIS rules and learn the user behavior. The S_i index helps in computing the merit target values on the factor of $O_{i,K}$. Here, ANFIS helps find the target value,

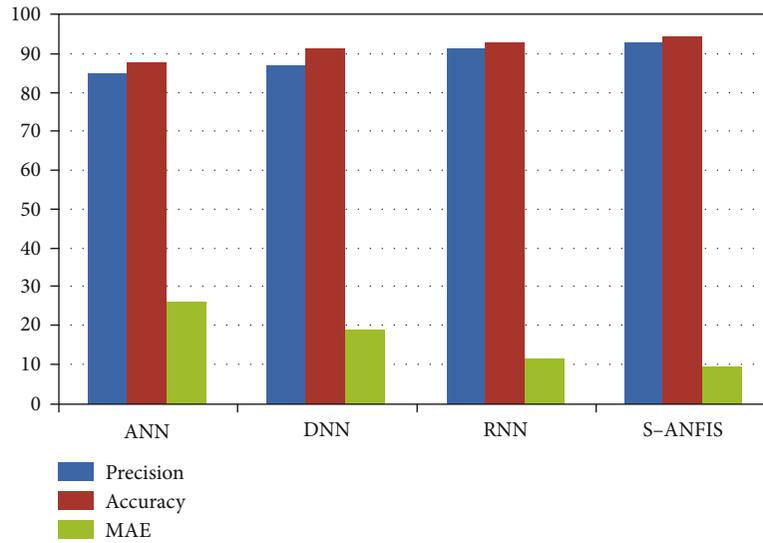


FIGURE 6: Training dataset ratio is 0.7 for PIKES.

TABLE 2: Outcome of various models implemented on the Hungarian blog dataset.

Model	Training ratio of 0.9			Training ratio of 0.8			Training ratio of 0.7		
	Precision	Accuracy	MAE	Precision	Accuracy	MAE	Precision	Accuracy	MAE
ANN	88.6	89.8	11.33	85.4	87.64	18.14	82.6	86.4	25.4
DNN	90.4	91.4	9	87.4	89.7	16.66	86.33	90.64	21.6
RNN	92	93.4	8.2	90.6	92.2	11.4	91	92.6	13.66
S-ANFIS	94.2	95.2	4.6	94	94.8	7.2	93.11	94.14	8.66

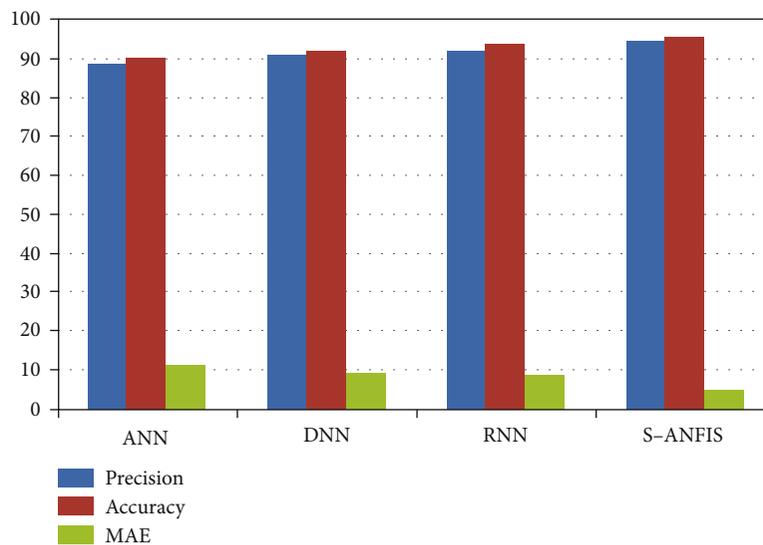


FIGURE 7: Training dataset ratio is 0.9 for Hungarian blog.

as a regression algorithm uses the learning capability of ANN; and the similarity index mapped with the inference rules, which results in the target value.

To inculcate all the modules in a single proposed model, the workflow will be shown as Figure 3.

The experiments are conducted on two different datasets for the demonstration purpose: first is PIKES (available at <http://pikes.fbk.eu/ke4ir.html>), and second is the blog data-

set which is UCI public data named as Hungarian blog. PIKES dataset majorly uses information retrieval and knowledge extraction techniques. It was published on May 29 to June 2, 2016. In this dataset, the text queries are stored, and ranking and query are appropriately defined. Finally, the aggregates.csv is the main file that has all the questions taken in this paper. Another dataset taken is the Hungarian blog dataset by UCI public data. This dataset has taken six

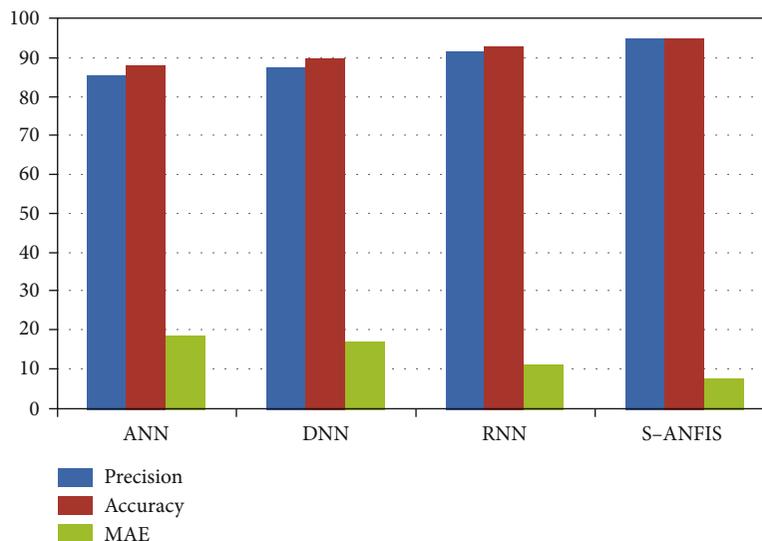


FIGURE 8: Training dataset ratio is 0.8 for Hungarian blog.

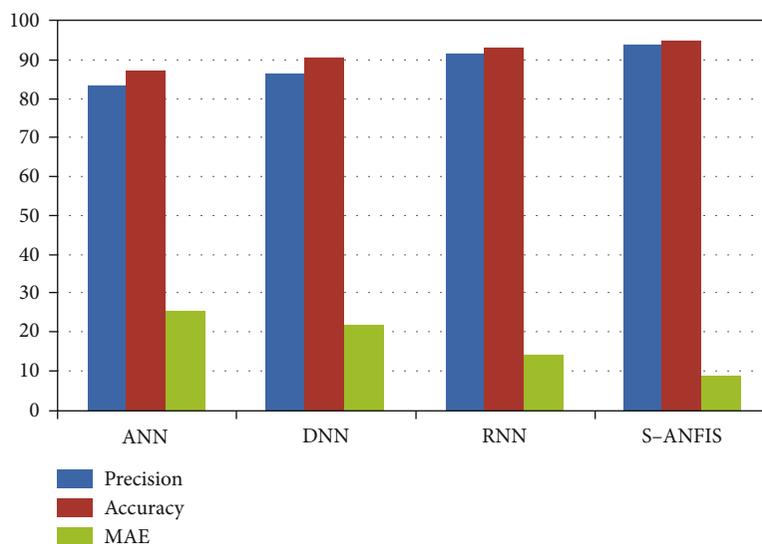


FIGURE 9: Training dataset ratio is 0.7 for Hungarian blog.

primary attributes and was used to check the performance of the proposed algorithm. It is a multivariate dataset without any missing values. The Hungarian blog dataset main fields are education; political caprice; topics; local media turnover; and local, political, and social spaces. The implementation and the results are shown in the next section.

4. Implementation and Results

The proposed model adopts the S-ANFIS model for the recommendations, and raw data of blog posts is gone through the content curation algorithm to extract the relevant key terms concerning blog posts. For the demonstration purpose, the experiments are conducted on two different datasets: for demonstration purpose, first is PIKES, an information retrieval dataset (<http://pikes.fbk.eu/ke4ir.html>), and the second is the blog dataset which is UCI pub-

lic data named as Hungarian blog. The algorithms are implemented using Python with Anaconda IDE. The experiments are carried out on a windows machine with an Intel i5 11th Generation Processor on 2.2 GHz having 16 GB RAM. The datasets are used on the three-proportion ratio: training data ratio: 0.9, 0.8, and 0.7. In addition, three hidden layers are considered for ANN, DNN, and CNN, the 1000 iterations are fixed, and the learning rate is 0.001.

4.1. Performance Metrics. Accuracy and precision are the performance parameters taken in the experimental results for the performance analysis.

- (i) Precision value
- (ii) Accuracy value
- (iii) Mean absolute error

4.2. Performance Analysis. The performance analysis of S-ANFIS in comparison to existing approaches is shown below. Among all methods, S-ANFIS gives the most relevant results in both datasets in all aspects. Table 1 shows the analysis of precision, accuracy, and mean absolute error matrices on ANN, DNN, RNN, and S-ANFIS models. Figures 4–6 show the results on discrete training dataset ratio for 1000 iterations on PIKES dataset.

Table 2 shows the analysis of precision, accuracy, and mean absolute error matrices on ANN, DNN, RNN, and S-ANFIS models. Figures 7–9 show the results on discrete training dataset ratio for 1000 iterations on Hungarian blog dataset.

5. Conclusion

This paper proposed an adaptive and intelligent content curation algorithm based on the hybrid model named S-ANFIS. The proposed algorithm was applied on two different datasets for demonstration purposes in this paper. Based on the implemented results, it is clear that the proposed approach can be used for actual web applications like Facebook, Twitter, and Instagram. The scope is not restricted to blogs and microblogging sites. It can also be applied to other domains like e-commerce sites, hotel booking sites, and movie review sites. Blogs are the best way to reflect the inner personality of the user's mind. So the examples of blog posts are taken to implement the algorithms and various models. In all the algorithms, S-ANFIS gives better results than ANN, DNN, and RNN when the accuracy, precision, and mean absolute error metrics are taken. In the future, ensuing algorithms might improve the existing results shown in the paper. The will also try to find other approaches to enhance the effects on the positive side with more performance metrics.

Data Availability

During the work, for demonstration purpose, the PIKES dataset is used which is available at <http://pikes.fbk.eu/ke4ir.html>.

Conflicts of Interest

During the work, no author has conflict of interest with any other author or organization.

References

- [1] Y. Jiugen, X. Ruonan, and W. Kewen, "Design and implementation of informal learning resources coconstruction based on microblog," in *2016 11th International Conference on Computer Science & Education (ICCSE)*, Nagoya, Japan, 2016.
- [2] C. Su, Y. Yu, M. Sui, and H. Zhang, "Friend recommendation algorithm based on user activity and social trust in LBSNs," in *2015 12th Web Information System and Application Conference (WISA)*, Jinan, China, 2015.
- [3] K. Harsh and K. B. Mohan, "A new approach to blog information searching and curating," in *Proceedings of the CSI 6th international conference on software engineering (CONSEG)*, pp. 372–377, Indore, India, September 2012.
- [4] H. Khatter, M. Trivedi, and B. Kalra, "An implementation of intelligent searching and curating technique on blog Web 2.0 tool," *International Journal of u-and e-Service, Science and Technology*, vol. 8, no. 6, pp. 45–54, 2015.
- [5] Y. Sun, X. Zhuang, K. Wei, X. Shan, and T. Jiang, "Using mobile services based on SNS to recommend who, how, and when to collaborate," in *2015 IEEE International Conference on Mobile Services*, New York, NY, USA, 2015.
- [6] M. A. Constanzo and S. Casas, "Usability evaluation of web support frameworks," in *2016 XLII Latin American Computing Conference (CLEI)*, Valparaiso, Chile, 2016.
- [7] H. Kaur and H. S. Pannu, "Blog response volume prediction using adaptive neuro fuzzy inference system," in *Proceedings of 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, Bengaluru, India, July 2018.
- [8] L.-A. K. Breuch, "Extended abstract: from research to practice: how insights from social web research inform website usability," in *2018 IEEE International Professional Communication Conference (ProComm)*, Toronto, ON, Canada, 2018.
- [9] K. Singh, R. Sandhu, and D. Kumar, "Comment volume prediction using neural networks and decision trees," in *IEEE UKSim-AMSS 17th International Conference on Computer Modelling and Simulation, UKSim2015 (UKSim2015)*, Cambridge, United Kingdom, 2015.
- [10] T. Harrington, R. Shenoy, M. Najork, and R. Panigrahy, "Patent: social network recommended content and recommending members for personalized search results," United States Patent, Patent No.: US 8-949, 232 B2, 2011.
- [11] H. Khatter and A. K. Ahlawat, "An intelligent personalized web blog searching technique using fuzzy-based feedback recurrent neural network," *Soft Computing*, vol. 24, no. 12, pp. 9321–9333, 2020.
- [12] X. Fu, M. Gao, Q. Liu, and H. Liu, "Group detection for specific topic on microblog," in *2016 13th Web Information Systems and Applications Conference (WISA)*, Wuhan, China, 2016.
- [13] J. Ding, Y. Dong, T. Gao, Z. Zhang, and Y. Liu, "Sentiment analysis of Chinese microblog based on classification and rich features," in *2016 13th Web Information Systems and Applications Conference (WISA)*, Wuhan, China, 2016.
- [14] X. Wang, Y. Jia, R. Chen, and B. Zhou, "Ranking user tags in microblogging website," in *2015 2nd International Conference on Information Science and Control Engineering*, Shanghai, China, 2015.
- [15] Y. Guo, M. Huang, and T. Lou, "A collaborative filtering algorithm of selecting neighbors based on user profiles and target item," in *2015 12th Web Information System and Application Conference (WISA)*, Jinan, China, 2015.
- [16] W. Gao and J. Wu, "A novel framework for service set recommendation in mashup creation," in *2017 IEEE International Conference on Web Services (ICWS)*, Honolulu, HI, USA, 2017.
- [17] Z. Borhanifard and B. Minaei-Bidgoli, "Recommender system for Persian blogs," in *2017 3th International Conference on Web Research (ICWR)*, Tehran, Iran, 2017.
- [18] N. A. B. M. Zulkefli and B. Bin Baharudin, "Hotel travel recommendation based on blog information," in *2015 International Symposium on Mathematical Sciences and Computing Research (iSMSC)*, Ipoh, Malaysia, 2015.

- [19] Y. Zhang, Y. Qian, and Y. Wang, "A recommendation algorithm based on dynamic user preference and service quality," in *2018 IEEE International Conference on Web Services (ICWS)*, San Francisco, CA, USA, 2018.
- [20] K. Harsh and A. K. Ahlawat, "An algorithmic approach for recommendation systems for web blogs and microblogs," *Journal of Xi'an Shiyou University*, vol. 16, no. 9, pp. 347–350, 2020.
- [21] H. Khatter, N. Goel, N. Gupta, and M. Gulati, "Movie recommendation system using cosine similarity with sentiment analysis," in *Third International Conference on Inventive Research in Computing Applications (ICIRCA)*, pp. 597–603, Coimbatore, India, 2021.
- [22] H. Khatter, S. Arif, U. Singh, S. Mathur, and S. Jain, "Product recommendation system for e-commerce using collaborative filtering and textual clustering," in *Third International Conference on Inventive Research in Computing Applications (ICIRCA)*, pp. 612–618, Coimbatore, India, 2021.
- [23] H. Khatter and A. K. Ahlawat, "Content curation algorithm on blog posts using hybrid computing," *Multimedia Tools and Applications*, vol. 81, no. 6, pp. 7589–7609, 2022.
- [24] H. Khatter and A. Ahlawat, "Web blog content curation using fuzzy-related capsule network-based auto encoder," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 36, no. 1, pp. 1–30, 2022.
- [25] Z. Ping, "Research on personalized blog information retrieval," in *Proceedings of the international conference on web information systems and mining (WISM) 2010*, pp. 289–292, Sanya, China, 2010.
- [26] H. Khatter and A. Kumar Ahlawat, "Analysis of content curation algorithms on personalized web searching," in *Proceedings of the International Conference on Innovative Computing & Communications (ICICC) 2020*, New Delhi, India, February 2020.
- [27] J. Rajeswari and S. Hariharan, "Personalized search recommender system: state of art, experimental results, and investigations," *Engineering*, vol. 6, no. 3, pp. 1–8, 2016.