CIWPR: A Strategic Framework for Collective Intelligence Encompassment for Web Page Recommendation

Manoj Kumar H S[1], Gerard Deepak[2], Santhanavijayan A[3]
[1,2,3]Department of Computer Science and Engineering
[1]PES University, Bengaluru, India
[2]Manipal Institute of Technology Bengaluru, Manipal Academy of Higher Education, Manipal, India
[3]National Institute of Technology, Tiruchirappalli, India
[2]gerard.deepak.christuni@gmail.com

Abstract. Web page recommendation is one of the vital strategies in the era of Web 3.0 due to the exponential increase of the contents over the World Wide Web (WWW). In this paper CIPWR framework for recommendation of web pages which incorporates upper ontology generation, ontology alignment using the Lin similarity with an appropriate threshold and SynSet generation has been put forth. The framework integrates two topic models or encopresis the concept of Bi-topic modelling with two distinct topic models namely the Latent Dirchlet Allocation (LDA) and Latent Semantic Indexing (LSI) for lateral enrichment of topics and also the NELL and YAGO knowledge stores are integrated for Query word enrichment which is further used for feature selection and a Random forest classifier which is a strong machine learning driven feature controlled classifier has been added for the classification of the dataset and semantic similarity is computed by amalgamating cosine similarity and Jaccard similarity with the differential thresholds in order to yield the best in class results of 97.81% of average accuracy, 96.18% of average precision, 99.43% of average recall, 97.78% of average F-measure and the lowest FDR of 0.04 has been accomplished by the proposed CIPWR framework.

Keywords: Collective Intelligence, Semantic Similarity, Web Page Recommendation, Web Page Retrieval, Web Search.

1. Introduction

Web page recommendation has increased its popularity in the recent years. It shows links to related stories or most visited websites. But there are many issues faced in developing a very effective and user-friendly web page recommendation system. Although there are a lot of information available to the users, there is a problem in finding the required information in the right time from large information source. This leads to information explosion. This is very crucial in e-commerce websites as they can lose their customers very easily. Although there are many other tools or approaches available in recommending web pages, a Semantic Method for recommending web pages is the need of the hour to comply to the Web 3.0 standards. Semantic Web is also called as Web 3.0 or also Data Centric Web and it is an expansion of the World Wide Web. Due to digitization, explosion of information on the web and lack of models in the current era which can support the Web 3.0 for recommendation of web pages there is a need for new framework which can help to solve this problem. So, this motivates to propose a more knowledge centric semantically inclined and semantically driven frameworks for recommending web pages. Also, the diversification the web pages triggers the need for a new framework which is more semantically inclined and knowledge centered. The accuracy of the all the baseline models is very less when compared to the proposed framework.

The proposed CIPWR framework has the feature of query word enrichment by using Bi-topic modelling with Latent Dirchlet Allocation (LDA) and Latent Semantic Indexing (LSI). The integration of NELL and YAGO knowledge stores for further enrichment. Upper ontology generation and ontology alignment with the generated knowledge and the entities from the knowledge stores for query word enrichment. The most important is the feature selection driven classification of dataset by subjecting it to strong Machine learning Random Forest classifier and SynSet generation and semantic similarity computation with two different measures i.e., Jaccard and Cosine similarity with differential thresholds for overall ranking and recommending of the web page. The average recall, average precision, average accuracy, F-measure are increased and the FDR value is decreased in comparison to the base line models.

2. Related Works

T. Gopalakrishnan et al. [1] have proposed a framework for recommending web pages with reduced time incorporating variable order Markov model with prediction accuracy. Suruchi Chawla [2] has proposed a method for searching the web which is according to web page communities' recommendation for personalized web search (PWS). The maximum flow method with HITS is used to create web page communities by clustering web pages with similar content that have been visited. Laxmi Rajani et al. [3] have proposed a model which makes web page recommendation using the Cumulative Page Weight (CP Weight) by adding the Google's PageRank, average time spent in that website and many more factors. Piyanuch Chaipornkaew et al. [4] have proposed a model for recommendation employing three machine learning models namely the TF-IDF, K-means and Apriori algorithms. The TF-IDF model was used for vectorization of words from webpage headings. K-means was utilized for clustering webpage headings and Apriori algorithm was used for finding association between the clustered webpages. The models employed to assess the accuracy were the KNN, Decision Tree, and Multi-Layer Perceptron respectively. To assess the prediction accuracy, KNN, Decision Trees, and Multi-Layer Perceptrons were used.

Satyaveer Singh et al. [5] have developed a semi-automatic ontology generation mechanism based on web mining methods, called a web page recommendation system. With the use of any web usage mining approach, a user's web data may be utilized to extract the user's behaviour. The meaning of web pages may be expressed using ontology, a form of knowledge representation. Xiaoyan Cai et al. [6] have created a deep network representation model that combines vertex content information with network structural information using generative adversarial networks to represent diverse vertices in a continuous and shared vector space. Deepak Surya et al. [7] have developed a method for mining web pages to retrieve current terms from user browsing history. WordNet is utilised to synonymize the extracted data, while Ant Colony Optimisation is employed to find the shortest path through graphs. Classification methods such as Random Forest classifier are also used. Gurunameh Singh Chhatwal et al. [8] have suggested a technique that takes into account the framework for semantic matching, entities that are enhanced by the creation of Resource Description Frameworks, and the inclusion of background information from the cloud of Linked Open Data. Entities from the Twitter API are also integrated, further promoting social awareness.

S. Anusuya et al. [9] have a model for a BAO+WSVM method for web page recommendation that combines Bat Agent Optimizations (BAOs) with Weighted Support Vector Machines (WSVMs). The user-system interface is formed by the user-bat agent, which gives users an interactive environment. By using Domain Ontology Web Language (DOWLs), which uses the website's structure to build the ontology, semantic extractions are performed. Naresh Kumar et al. [10] have utilised REST APIs to crawl the web. Additionally, they have combined collaborative filtering with content-based filtering to create a hybrid system for recommendations.Pradeep Bedi et al. [11] have proposed a model which classifies related sites by semantic matching of the quest route. Duc-Hieu Tran [12] has proposed a technique for semantic indexing and retrieval of web pages. A content extraction model is applied to extract the content of the bookmarked web page. This is semantically indexed and stored in URL. Semantic indexing and retrieval enable users to retrieve a web page that includes content semantically linked to the search query. Hiteshwar Kumar Azad et al. [13] three weighting models based on k-nearest neighbour (kNN) based cosine similarity, TF-IDF, and co-relation score have been put forth as a query expansion strategy. In [14-18], several models which were correlative to the proposed framework are depicted.

3. Proposed System Architecture

The structure of the framework for web page recommendations has been depicted in the Fig. 1. This is a knowledge centric ontology driven web page driven model that integrates a hybrid topic modelling scheme for entity enrichment and uses a Random Forest classifier. This approach involves several steps. The user query, which goes through pre-processing, is the first input. The steps involved in query pre-processing include tokenization, lemmatization, named entity recognition (NER), and stop word elimination. Tokenization is the process of disassembling the texts into concise units called tokens that may be utilised in a database or internal system without bringing them into scope. Lemmatization produces root words or synonyms of inflected words so that they are analysed as one term or in simple words it refers to the process of finding the normalized form of the word. The words that often appear are eliminated during the stop word removal procedure since they bring no value to the analysis, lengthen processing time, and increase the dimension of the feature set. In general removing stop words significantly increases the performance. The Python package used for stop word removal is called NLTK. Identification and classification of the data or item into predetermined categories is known as NER. So, at the end of the pre-processing stage, the individual query words i.e., Q_w are obtained or yielded.

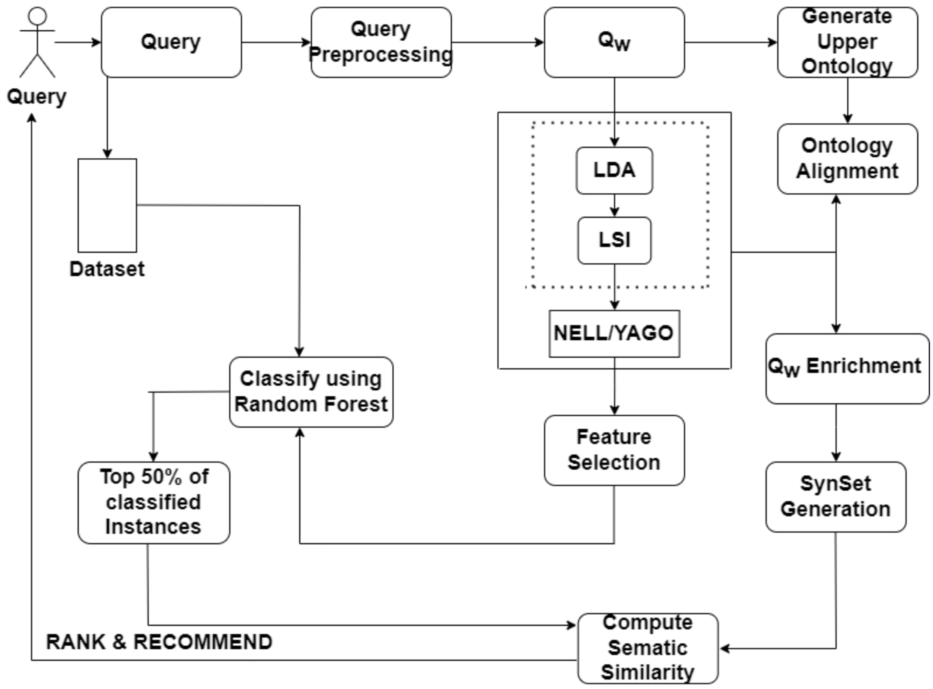

Fig. 1. Proposed System Architecture

Once the query words are yielded it is subjected to topic modelling. The speciality of this model is that the topic modelling is done using two distinct individual approaches namely Latent Dirichlet Allocation (LDA) and Latent Semantic Indexing (LSI). The advantage of using this double topic modelling is that they uncover the hidden and yet relevant terminologies related to the query words from the external documentary corpus or the World Wide Web. This is how LDA and LSI are used and these topic models enrich the entities. However, entity enrichment is further done by leveraging to knowledge sources like NELL / YAGO. These are two distinct knowledge sources. Here the topics yielded by the query words are submitted to NELL or YAGO via the API and the entities are yielded which is used for feature selection. Feature selection takes place by using Shannon Computational Entropy. The selected features are passed to the Random Forest classifier for classifying the dataset.

Random Forest classifier is also called as a Supervised Learning algorithm as it is used to train algorithms and predict the output the outputs accurately. Random Forest is an ensemble machine learning model or algorithm that is used widely in classification and regression and many more tasks. Many decision trees are built and trained for the above task. The output class is the class that most decision trees choose when the dataset is categorized using the Random Forest classifier in the suggested CIWPR framework. Based on the Bagging technique, the Random Forest model randomly chooses a subset of characteristics from the training set and bootstrap samples for each decision tree.

The top 50% of the outcome of the Random Forest classifier instances are yielded under each class in order to target only the relevant entities. Subsequently the query words are used to generate the Upper Ontology. Ontology is a formal specification that provides reusable and sharable knowledge representation. Upper ontology is generated using OntoCollab as a tool. Upper ontologies are used only up to 4 levels excluding the individuals and the ontology alignment is performed between the ontologies aligned and the entities enriched based on LDA, LSI.

Lesk Semantic Indexing (LSI) is based on the principle that the words that are used in similar contexts have the same meaning and hence this makes it easier to extract the conceptual context from the text by establishing connections between words that have the same meaning. Latent Dirichlet Allocation (LDA) is used for discovering latent semantic topics in text documents. Hence also called as Topic Modelling. It is just not used in word or words similarity.

Cosine Similarity is based on vector space model. Cosine Similarity is a measure of how similar two data points are there in the plane. It measures the angle of cosine between two objects. It can be found out by finding the dot product between the two identities.The cosine similarity is depicted in Equation (1).

$$\left(\overrightarrow{t_a}, \overrightarrow{t_b}\right) = \frac{\overrightarrow{t_a} \cdot \overrightarrow{t_b}}{|\overrightarrow{t_a}| x |\overrightarrow{t_b}|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \sqrt{\sum_{i=1}^{n} B_i^2}} \tag{1}$$

Equation (1) depicts the equation of the Cosine Similarity where $\overrightarrow{t_a}$ and $\overrightarrow{t_b}$ denotes the attributes of two vectors A and B. A_i and B_i denotes the attributes of the vectors A and B.One of the similarity metrics used to assess the degree of similarity between two items, such as two texts, is the Jaccard similarity. Additionally, it may be applied to determine whether any two asymmetric binary vectors or two sets are comparable to one another. It is also known as Jaccard Index, Jaccard Similarity and Jaccard Distance. Equation (2) depicts the Jaccard Similarity where A and B are two data sets. It is also written as J (A, B) = (Total number of observations across the two sets) (number of observations in either set).

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \tag{2}$$

4. Implementation, Results and Performance Evaluation

The implementation part was carried out using Python 3.10.6 using Google's Collaboratory as the Integrated Development Engine or the IDE. Python's natural language toolkit for used for pre-processing as well as the natural language processing tasks to accomplish the pre-processing and natural language processing tasks. Upper ontology generation was carried out using OntoCollab as a tool. The PC has an Intel Core i7 CPU with 32 GB of RAM and an 8x 2.5 GHz-4.8 GHz clock speed. Experimentations were performed for a single large dataset which was formulated by integrating three distinct individual datasets namely the CrowdFlower dataset, Raw URL Dataset and the TURL Dataset. All these three individual datasets namely CrowdFlower dataset, Raw URL Dataset and the TURL Dataset were annotated wherever it was necessary at first. The performance of the proposed CIPWR framework is evaluated using average Recall, average Accuracy, average Precision, F-measure percentages, and False Discovery Rate (FDR) as preferred potential standard metrics. Precision, Accuracy, Recall, F-measure quantify the relevance of the result whereas the false discovery rate quantifies he number of false positives captured or furnished by the framework which is nothing but the error rate. Synthesized error rate generated by the CIPWR framework. For evaluating the performance of the CIPWR framework it is baselined with WPMMR, IWSPWC, WPOC, RMWC and COWPR frameworks respectively. Both the baseline frameworks and the proposed CIPWR framework were evaluated in the exact same environment in order to benchmark and compare the performances yielded for the exact same data set or the exact same number of queries as that for the CIPWR framework. In order to evaluate the results for the CIPWR, the experiments were conducted on 4814 queries for which the ground truth was carried out for a period of 7 months. 63 candidate users were given the set of queries, not all of them got all the queries but each of them got at least 20-40 queries every month and periodically they were asked to give their top 10 searches in terms of the returned query face sets as well as the web pages and the key words on the web pages from their favorite search engines like Google, Bing, Brave, Safari etc., where they gave their ground truth.

Table 1. Comparison of Performance of the proposed CIPWR with other approaches

Model	**Average Precision %**	**Average Recall %**	**Accuracy %**	**F-Measure %**	**FDR**
WPMMR [1]	84.12	86.18	85.15	85.14	0.16
IWSPWC [2]	87.22	89.39	88.31	88.29	0.13
WPOC [3]	87.23	90.02	88.63	88.6	0.13
RMWC [4]	89.36	90.14	89.75	89.75	0.11
COWPR [5]	92.19	94.63	93.41	93.39	0.08
Proposed CIWPR	96.18	99.43	97.81	97.78	0.04

It is indicative from Table 1 that CIPWR yields the highest average recall of 99.43%, highest average precision of 96.18%, highest accuracy of 97.81%, highest F-measure of 97.78% and lowest FDR of 0.04. The reason why the proposed CIPWR gives the highest recall, accuracy, precision, F-measure, and lowest FDR is mainly because of the fact that it is strategically driven by queries and auxiliary knowledge. So, it is a knowledge driven framework where pre-processors query is subjected to upper ontology generation which is generally core concepts, sub-concepts generation along with instance population. Apart from this the query words are staged strategically and enriched by subjecting it to Latent Dirchlet Allocation (LDA), Latent Semantic

Indexing (LSI) and also NELL and YAGO knowledge sources which are two heterogenous knowledge sources and made use in order to anchor entities. The LDA and LSI both are probabilistic generative topic models which yields uncovered topics from surrounding web corpora or the surrounding documents. In this case the web scraped web corpora is being anchored. The terms which are relevant and yield undiscovered to one of the query words are being anchored by means of LDA, LSI and the NELL and YAGO knowledge sources which enrich and populate entities which contain community contributed or community verified linked open data which strategically, exponentially increases the density of the entities by growing these entities. Subjectively the upper ontology is aligned with the grown entities which are discovered by LDA, LSI, NELL and YAGO models respectively for query enrichment. Apart from this the enriched queries are further subjected to SynSet generation using the WordNet 3.0 and most importantly the data set is classified using a strong feature-controlled Machine Learning classifier. The reason for not using a Deep Learning classifier here is because they work on the principle of auto handcrafted feature selection and auto handcrafted feature of Deep Learning framework can over learn even the outliers and therefore reduce the subjective precision, recall, accuracy, F-measure. Henceforth the feature-controlled machine learning classifier which is strong in Random Forest classifier is essential. Apart from this as the proposed model arguments the auxiliary knowledge at a very high rate and there is mechanism for regulation by means of Ontology Alignment using the LIN Similarity with the selective threshold and also amalgamation of feature-controlled Random Forest classifier and Semantic similarity using heterogenous measures with differential thresholds like the Jaccard and the Cosine Similarity ensures the proposed CIWPR model is a Collective Intelligence framework for recommending the web pages and this is the best approach the outperforms the baseline models.From the Table 1 it is also indicative that the WPMMR model has an overall precision percentage of 84.12%, overall recall percentage of 86.18%, overall accuracy of 85.15%, overall F-measure percentage of 85.14% and an FDR of 0.16. The IWSPWC model yields an overall precision percentage of 87.22%, overall recall percentage of 89.39%, overall accuracy of 88.31%, overall F-measure percentage of 88.29% and an FDR of 0.13. Similarly, WPOC model yields an overall precision percentage of 87.23%, overall recall percentage of 90.02%, overall accuracy of 88.63%, overall F-measure percentage of 88.6% and an FDR of 0.13. The RMWC model gives has an overall precision percentage of 89.36%, overall recall percentage of 90.14%, overall accuracy of 89.75%, overall F-measure percentage of 89.75% and an FDR of 0.11 and the COWPR model furnishes an overall precision percentage of 92.19%, overall recall percentage of 94.63%, overall accuracy of 93.41%, overall F-measure percentage of 93.39% and an FDR of 0.08.The reason why WPMMR model does not perform as good compared to the proposed CIWPR mode is because although WPMMR model is a web page prediction method, it’s a very traditional method which uses variable order Markov model and Fuzzy C-means clustering is applied on to the web log. However, there is no strong regulatory mechanism in terms of learning, in terms of aggregating auxiliary knowledge and regulating the auxiliary knowledge. There is absolutely zero percentage of auxiliary knowledge augmented with the model. So, prediction using a traditional statistical model and Fuzzy C-means clustering does not work when the web data is overcrowded and is highly cohesively condensed as in Web 3.0. It is also that the WPMMR model is imprecise in terms of its prediction and recall, accuracy, F-measure. So, basically it can never be scaled up for the Web 3.0 as it is based only on a traditional statistical Markov model wherein Web log alone is used and there is no auxiliary knowledge and no learning mechanism and Fuzzy C-means clustering fails strategically. Henceforth this model is not as accurate compared to the proposed model.

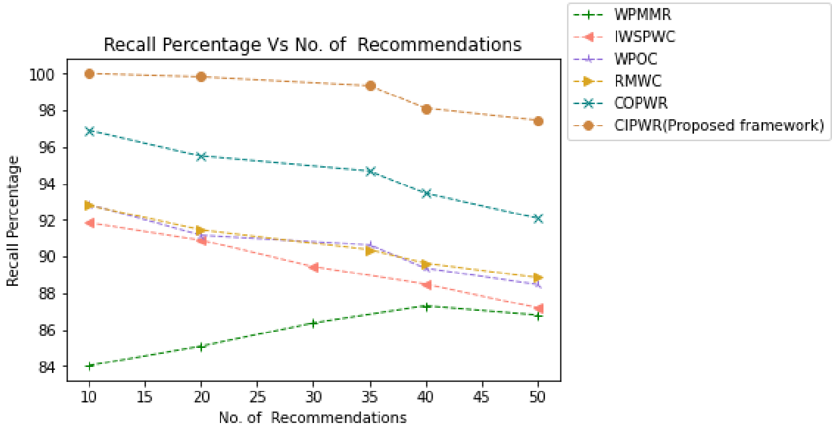

Fig. 2. Recall % vs Number of Recommendations Distribution Curve

The IWSPWC model which is a personalized web search based on web page communities also does not perform as excepted because it incorporates all the paradigm of hyperlink induced topic search which is HITS algorithm which is also based on clustering. Again, it is a community-based algorithm on selection of web pages based on the maximum number of hyperlinks by a community or group of community. So, based on the number of HITS the pathway is directed which is a very naïve and a traditional approach which basically doesn’t argument auxiliary knowledge. It does not work at all when the density of the data becomes very high as in Web 3.0. Although the WPOC model is a web page recommendation for organizational users using collaborative page weight it is a very naïve strategy which induces cumulative page weight over the traditional page rank algorithm where the hyperlinks based on the community rank pages will be discovered and this is suitable for Web 2.0. d.The reason why the RMWC framework does not perform as excepted when compared to the proposed model is because although RMWC model is based on user behaviours which incorporates K-means clustering, Apriori algorithms and TF-IDF. It is a hybrid strategy. It does not have a strong learning mechanism although they use KNN decision tree multilayer perceptron’s. But these mechanisms are not strong coherently because of the fact that they lack auxiliary knowledge and they depend only the data set. Classifying and learning from the data set itself has a lot of outlier-based learning because of lack of knowledge fed into the model owing to this reason although this model hybridizes TF-IDF, K-means clustering algorithms, KNN decision tree multilayer perceptron.

The reason COPWR model fails although it is a semantically driven, accumulating knowledge in the form of strategic effective ontology is due to the fact that ontologies are constructed statically and static knowledge is incorporated. So static knowledge although it yields perfectly relevant results. Availability of static knowledge in Web 3.0 is like finding needle in a haystack. It is absolutely impractical to use it. So thereby there is no regulatory mechanism for generations of ontologies or to regulate the knowledge or to extract ontologies from large amount of knowledge. Also, this lacks strong recommendation framework, strong comparative frameworks and lacks training based on classifiers. Owing to this reason this model which is effective ontology construction for web page recommendation does not perform as expected when compared to the proposed model.Owing to all these reasons the proposed CIPWR model which is collective intelligence driven framework which augments knowledge from several sources, grows knowledge, enriches knowledge, uses knowledge from secondary knowledge stores dynamically generates knowledge at every stage and have strong relevance, computation, strategies in terms of semantic similarity and step deviance, measures, accumulates ontologies which are generated and also dynamic and static knowledge is accumulates and apart from this has a strong machine learning feature selection model which is the random forest classifier. Hence from all these reasons the proposed model outperforms all the baseline models. Fig. 2. depicts the line graph of Number of Recommendations distribution Vs Precision curve for all the approaches. It is clear that the given I-DLMI model occupies the highest in the hierarchy.

6. Conclusions

Web page recommendation is one of the most important strategies in order to tackle the largely growing voluminous data on the web. This paper proposes CIPWR framework for recommending web pages which is knowledge centric and semantically inclined in order to accommodate the cohesive structure of the Web 3.0. So, the proposed CIPWR framework enriches the query words by subjecting it to two distinct heterogenous models namely LDA and LSI for topic modelling and lateral knowledge enrichment by means of NELL and YAGO knowledge stores along with generation of upper ontology with ontology alignment which enriches query enrichment and feature selection. The classification of dataset is achieved by using the Random Forest classifier which is a strong machine learning feature-controlled classifier. SynSet generation, ontology alignment as well as the semantic similarity computation with semantic similarity measures namely the Cosine and the Jaccard similarity with differential thresholds ensures high diversity of results and the best-in-class accuracy of 97.81%, precision of 96.18%, recall of 99.43%, F-measure of 97.78% with the lowest FDR of 0.04

References

1. Gopalakrishnan, T., Sengottuvelan, P., Bharathi, A., & Lokeshkumar, R. (2018). An approach to webpage prediction method using variable order Markov model in recommendation systems. Journal of Internet Technology, 19(2), 415-424.
2. Chawla, S. (2019). Intelligent web search system for personalised web search based on recommendation of web page communities. International Journal of Intelligent Systems Design and Computing, 3(1), 12-30.

3. Rajani, L., & Thakar, U. (2021). Webpage Recommendation for Organization Users via Collaborative Page Weight. Journal of Scientific Research, 65(1).
4. Chaipornkaew, P., & Banditwattanawong, T. (2021, May). A recommendation model based on user behaviors on commercial websites using TF-IDF, KMeans, and Apriori algorithms. In International Conference on Computing and Information Technology (pp. 55-65). Springer, Cham.
5. Singh, S., & Aswal, M. S. (2021). Constructing an effective ontology for web page recommendation. International Journal of Web Engineering and Technology, 16(2), 86-112.
6. Cai, X., Han, J., & Yang, L. (2018). Generative adversarial network based heterogeneous bibliographic network representation for personalized citation recommendation. In Proceedings of the AAAI conference on artificial intelligence (Vol. 32, No. 1).
7. Surya, D., Deepak, G., & Santhanavijayan, A. (2021). KSTAR: a knowledge based approach for socially relevant term aggregation for web page recommendation. In International Conf/ on Digital Technologies and Applications (pp. 555-564). Springer.
8. Chhatwal, G. S., & Deepak, G. (2022). IEESWPR: An Integrative Entity Enrichment Scheme for Socially Aware Web Page Recommendation. In Data Science and Security (pp. 239-249). Springer, Singapore.
9. Anusuya, S., & Mallika, N. M. Web page recommendation system using bat optimization and weighted support vector machine algorithm for health care service. International Journal of Health Sciences, (II), 5296-5317.
10. Kumar, N., Gupta, M., Sharma, D., & Ofori, I. (2022). Technical Job Recommendation System Using APIs and Web Crawling. Computational Intelligence and Neuroscience, 2022.
11. Bedi, P., Goyal, S. B., Rajawat, A. S., Shaw, R. N., & Ghosh, A. (2022). A framework for personalizing atypical web search sessions with concept-based user profiles using selective machine learning techniques. In Advanced Computing and Intelligent Technologies (pp. 279-291). Springer, Singapore.
12. Tran, D. H. (2022). Semantic Indexing and Retrieval of Web Pages in a Web Browser.
13. Azad, H. K., Deepak, A., Chakraborty, C., & Abhishek, K. (2022). Improving query expansion using pseudo-relevant web knowledge for information retrieval. Pattern Recognition Letters, 158, 148-156.
14. Deepak, G., Priyadarshini, J. S., & Babu, M. H. (2016, October). A differential semantic algorithm for query relevant web page recommendation. In IEEE International Conf. on Advances in Computer Applications (ICACA) (pp. 44-49).
15. Roopak, N., & Deepak, G. (2021, November). OntoKnowNHS: ontology driven knowledge centric novel hybridised semantic scheme for image recommendation using knowledge graph. In Iberoamerican Knowledge Graphs and Semantic Web Conference (pp. 138-152). Springer, Cham.
16. Ojha, R., & Deepak, G. (2021, November). Metadata driven semantically aware medical query expansion. In Iberoamerican Knowledge Graphs and Semantic Web Conference (pp. 223-233). Springer, Cham.
17. Rithish, H., Deepak, G., & Santhanavijayan, A. (2021, January). Automated assessment of question quality on online community forums. In International Conference on Digital Technologies and Applications (pp. 791-800). Springer, Cham.
18. Yethindra, D. N., & Deepak, G. (2021, September). A semantic approach for fashion recommendation using logistic regression and ontologies. In 2021 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES) (pp. 1-6). IEEE.