Journal of Theoretical and Applied Information Technology
15th October 2022. Vol.100. No 19

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

INTILLEGENT WATER DROP ALGORITHM BASED WEB PAGE RECOMMENDATION USING LINEAR REGRESSION LEARNING

PAVITHRA B [1], NIRANJANAMURTHY M [2]

[1] Research Scholar, Department of Computer Applications, M S Ramaiah Institute of Technology (Affiliated to Visvesvaraya Technological University, Karnataka), Bangalore 560054, Karnataka, India.

[2] Department of Computer Applications, M S Ramaiah Institute of Technology (Affiliated to Visvesvaraya Technological University, Karnataka), Bangalore 560054, Karnataka, India.

Email: [1] pavithra247@gmail.com , [2]niruhsd@gmail.com

ABSTRACT

Analysis of website visitor for improving the performance of user experience on web portal is primary requirement in this online world. Hence number of web-based feature was used for the study of user behavior and recommend page accordingly. The problem is we don't have any sophisticated method to predict the pages with greater precision value in association with the faster accessing speed of the webpages this research work presents the utilization of weblog feature for analysis and pattern learning. Proposed model extract association rule forms the weblog with text pattern format on the web page. Here we have considered the input data set from a live website's weblogs called as project tunnel.com. As an output the model gives the probability of pages predicted in a sequence manner, thus increasing the precision values of predicted pages and also it increases the chromosome page quality involving a minimal value of fitness function with the regression value. The proposed work has concluded with having increasing precision value with minimal fitness value of a precision value when to compare with existing methods like Genetic feed forward association and Particle swam optimization. Association rule and text URL patterns were used for learning multi-Regression probability generation. Once the regression value is set than Intelligent Water Drop Algorithm will be applied for predicting or recommending the next page as per user previous set of pages, where IWD algorithm increases the accuracy of prediction by utilizing the regression values in fitness function. Live websites weblogs were analyzed for the experiment and comparison of proposed WPPIWD (Web Page Prediction Intelligent Water Drop) has improved the work efficiency when compared to other existing methods. Results shows that proposed intelligent water drop based Web Page Prediction Model (IWD) for next page prediction system has improved various evaluation parameters like precision, coverage, metric.

Keywords: *Information Extraction, Neural Network, Recommendation, Regression, Weblog.*

1. INTRODUCTION

Web mining is treated as a data mining application, it deals with the activity deprived on rules obtained from web analytics. It finds user login behavior or login patterns from the web server registry using data mining techniques to predict user browsing trends. Knowing about the user login patterns which deals in helping the webmasters customize website content to extend a user's exploring proficiency [1].

Today, issues of modeling and predicting behavior of the user on websites has impacted on search interest because it can be applied in web cache performance improvement, web page recommendations, engine improvement research, today, issues of modeling and predicting behavior of the user on websites has impacted on search interest because it can be applied in web cache performance improvement, web page recommendations, engine improvement research

Modeling, and browsing personalization. Experience. Analysis of the web observing interesting facts of the customer using it, is the most vital and interesting research topics in mining the use of the web. If users 'interests can be spontaneously identified from users' web registration data, after that it can be implemented to recommend information and marketing that is useful to both users and website developers. [3, 4]. In the bibliographic research column, most of the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

implementation used to recommend websites were accomplished with a property of parsing the text in the center with the methodologies considered to gear towards the framework of mining the texts phrases.

Let say, in the current context, the websites is getting multidisciplinary description along with page info that contain images, video, audio etc. Recently, multimedia description website recommendation procedures have been introduced, and there is a long way to go towards the end. The main motivation for this work is the introduction of an imparted agent oriented structure with an optimization methodology. Many imparted agent oriented website recommendation algorithms using semantic information have been introduced in recent decades, but in previous work searching for agent information has become slower. To solve this problem, an optimization methodology has been discussed in the article

Many researchers in the area of web usage mining have conducted research to analyze weblog data and discover common weblog data patterns. The foremost important objective of studying the user's visitor logs is to analyze the users who are mainly worried about the final outcome of the mining results.

By evaluating the user access to the content over time, frequency, etc., you are changing the structure and layout the website to expect more users to interact and customers to receive better service. The analysis of user behavior has now been a vital research topic and access point. The work of this process mainly explores the web log mining technology in the analysis of user behavior and creates the user interesting behavior and finally the information of the user's attention and, lastly, attracts the user's interest in webpages

In the bibliographic research column, most of the implementation used to recommend websites in centralized text oriented and then methodologies were geared towards the structure of the texting. Moreover, in the current context, the web is getting multimedia description with pages that contain images, audios, videos, etc. Recently, multimedia description website recommendation procedures have been introduced, and there is a long way to go towards the end. The main motivation for this work is the introduction of an agent-based framework with an optimization algorithm, many imparted agent oriented website recommendation algorithms using semantic information have been introduced in recent decades, but in previous work searching for agent information has become slower. To Look into this issues, an optimization methodology has been discussed in the article

In continuation further, this paper was implemented into new subtypes. Related section have related work content where next page prediction researcher work were explained. After this proposed work WPPIWD methodology was detailed with block diagram and algorithm. Fourth section provide experimental value comparison with other existing methods. Finally whole work is concluding with improvement percentage.

2. RELATED WORK

In [7] propose four models based on LSTM. Initially, non-Interact LSTM, a visibility prediction value is issued at each step. The values which is given for very iteration it involves the value about the user, page, depth, and context. The fore coming parameters characteristics are granted by make use of the keying levels. The second model, LSTM Interact, progresses these value by taking into account user interaction, page, and depth; the architecture involves duplicating the surrounded values in prior sending the values to LSTM introverts. The next model, BLSTM Interact, takes into account the importance that scroll through pages frequently; this bidirectional method takes into account the dwell time sequence in two directions of travel. The proposed next model, RED BLSTM Interact, will make use of the residual connections and vailing and unveiling structure prolonged with the BLSTM hyper values layers to work positively with the piled up LSTM layers and overcome the problem of gradient masking and data irrelevant.

In [8] introduced an instance automated system to address the scalability issues which is a common issues to most of the data mining techniques and provide users with proper recommendations oriented on their current browsing sequence. They used the nearest neighbor rating and were able to provide very precise filtering based supportive idea on the interesting current needs rather than the interested human of previous visits. Other data mining techniques may be used in the future.

In [9] assumed in the recommendation process using fuzzy c mean grouping for the interests of many users. C mean Fuzzy grouping maps data items to multiple clusters or to one cluster.

This method involves a vital work for new websites because website owners can keep track of who will buy hence forth and suggest recommendations. The adding of privacy also certainty of the privacy holds on system is required. Erman Sen, Hakki Toroslu, Pinar Karagoz [6] has proposed the summarizing of pages action by make use of grouping, which is spastically message of an extended versions [2]. It viewed semantics likely to be a group of concert related to the semantics of this page and used a concept-based semantic similarity method for grouping. This semantic alikeness improves precision. You have a set of manually likely to be a group of concert related to the semantics mythology was developed logically, so computerization is must in this phase. You created and saved all assemblies for guessing, but it is not practical to save all sessions, so the comitia involved in picking up the sessions used in predicting should be provoked. We have expanded the work of [6] by making use of the contents of user for the coverage by dissolving the implementation of discarding procedure

Efforts have been imposed to the literature to address the drawbacks of using advertisement space for long-tail questions, particularly for proposed search. Many of the work to get advertisement space for final searches is done through the search extension. In [13] a method was proposed to differentiate queue search queries using a model based on the implication of concepts. Organic based images were used as a backup mechanism to yield the hierarchy-based conceptual model. Another method [12] broadens search questions by implementing more functionality, using the categorization of web pages accessed and prominent named property.

In [11], an approach has been proposed to classify the final search queries using a model based on the hierarchy of concepts. Organic based images were used as a default measure to learn the hierarchy-based conceptual model. Another method [11] extends search questions by implementing more functionality, including the category of web pages accessed and prominent named entities.

In [10] found that evaluating website functionality is a time-consuming and physically done task. Grants a mechanism that extends its favor of remote capacity knowing or predicting functionalities of Internet Supports client towards message on single independent dispatches the message details. In addition, the importance of tailored opportunity offers verifies the adaptability to discover certain events and also to take them into account involving anatomizing. The instruments offered supports the website by editing a proxy-based style and permits the attacker to take real individual actions and an optimal sequence of steps.

Nguyen. [14] Implemented a recommendation system that will propose and redirect the user choice of courses depending on her needs. The hybrid technique was used in Ontology Society to extract valuable information and provide accurate methodologies. These approach are beneficial for students to improve their knowledge and their level of satisfaction. The featured recommender systems will work best by removing the constraints from the basic individual recommender systems.

Problem statements that has been identified after reviewing states that many optimized algorithms which was proposed tends to be used for the aim of clustering datasets of an oversized scale, but due to low precision value and maximum utilization of the web indexes large scaled data was not perfected effectively, it is also notice that by using particle swam agent as one of the optimized technique it has a drawback of clearing out previously saved agent item predicted page details, if so then successive web page details will not be saved in the frequently visited cache there is a chance of missing out the frequently visited page agent and thus missed the accuracy of the prediction

3. PROPOSED METHODOLOGY

The solution for the above stated problem is given through this work, the proposed work WPPIWD embeds machine learning approach and try to get the frequently visited web pages without missing out the previous users history, within the accessible fitness value and low precision accuracy and also meeting some of the parameters which is discussed in detail further. The work explained in this section can be reflected in the fig. 1 which shows block diagram for training of machine learning model. Here log features were used for the working of the training and testing. Explanation of testing model were done in fig. 2 where proposed intelligent water drop algorithm was used for prediction of page s per visitor.

Weblog Pre-Processing

Input weblog dataset need to pre-process first, so unwanted information like operating system, browser information are remove. Here as

Journal of Theoretical and Applied Information Technology
15th October 2022. Vol.100. No 19

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

per pattern in weblog information were select or reject. This pre-processing select few information from the log first is IP address "222.205.248.120"second is date and time and third is visiting URL

URL Pattern

Pre-processed data URL were taken in this step here text present after last backslash '/' is taken as page "online-hotel-management-project-php". So whole weblog set of pages were filter and each unique page is identified by a number term as Page ID, so if N pages are present in the weblog than P1-N. Now terms are collect from the page name like: Now each filter page is [online, hotel, management, project, and php". Develop a matrix having dimension NxN where number of similar terms between pages were manage.

3.1 Association Rule:

Each unique IP address is consider as single user where user set of pages were identified as per date, time sequence. Here as per Page ID sequence of pages visit by one user is arrange, where high difference in time sequence is consider as separate sequence. So this give a set of webpage sequence visit by users, each set is arrange in row of Log matrix having dimension Lx1, where L is number of logs obtained from dataset.

Now Log matrix is pass in the Feed Forward counter proposed in [2]. Feed forward counter gives a set of rules in one scan of dataset. This set of dataset is further used for the machine learning.

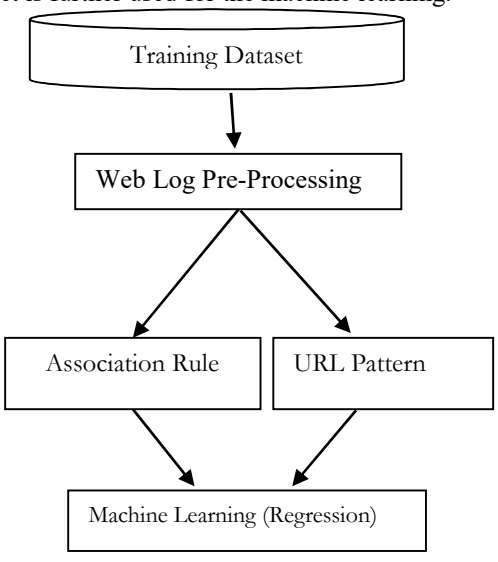

Fig. 1 Training of Web Page prediction model.

3.2 Regression

Learning from the URL pattern and association rules were done by multi regression. So similar set of association rule from left hand side were collect in a single bunch which act as input from association rule. While URL pattern value for the two page similarity act as another input value. This can be understand as let association rule P2, P9 →P1 and other rule is P2, P9 →P7 than page have similar left hand side rule confidence value Crule1, Crule2 were used. While URL pattern value for [{P2, P9}, {P9, P1}, {P2, P1}, {P2, P7}, {P9, P7}] is pass in form of URL pattern. So input matrix in regression have three row and two column (as per number of association rule).

$$\text{Input} = I \begin{matrix} C_{rule1x\{P2,P9\}} & C_{rule1x\{P9,P1\}} & C_{rule1x\{P2,P1\}} \\ C_{rule1x\{P2,P9\}} & C_{rule1x\{P9,P7\}} & C_{rule1x\{P2,P7\}} \end{matrix}$$

Hence output of this regression is intercept β1, β2, β3. So as per rules left hand side values are store in Association rule matrix R. This is learning of data from the input dataset.

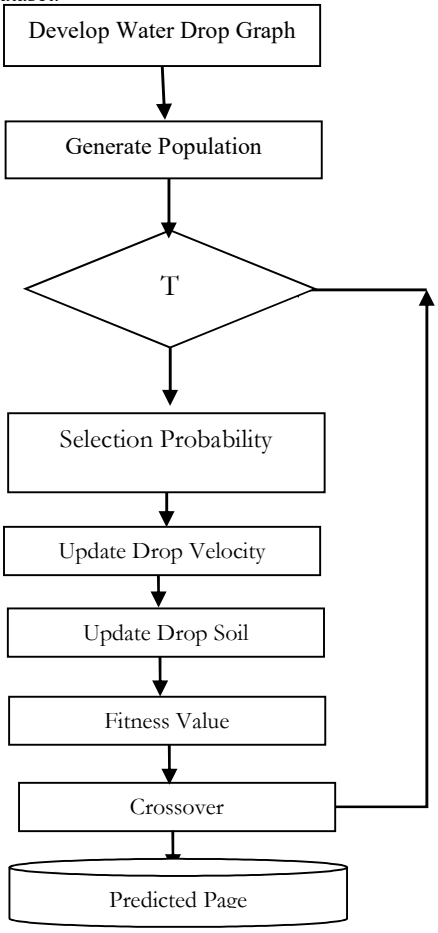

Fig. 2 Block diagram of proposed model.

Journal of Theoretical and Applied Information Technology
15th October 2022. Vol.100. No 19

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3.3 Testing Module

Input testing dataset have logs which need to preprocess as number of pages visit by the single user. Hence page sequence were generate which is pass in the intelligent water drop algorithm to predict next page.

WPPIWD (Web Page Prediction Intelligent Water Drop)

With respect to genetic algorithm 'DROP of WATER' tends to be as a node in a graph while association rule confidence value act as graph edge 'Soil'. So user move as per matching page content which was judge by feature difference [13]. So higher Soil value means higher lower chance to predict that page.

Graph of Water Drop

It terms to be a finalized graph with N number of nodes, where N is the collection of number of pages. Therefore, GWD is NxN. The weight value of the graph in the form of SOIL S was estimated for individual edge which is considered amongst the nodes as per the values to the left of the association rule. This value is given by equation 1.

$$Soil(i,j) = \frac{1}{C_{ruleij}} \text{ --Eq. 1}$$

Static and Dynamic Parameter

With regards to this process, some of constant will be initialized in the beginning of algorithm and its termed as updating parameters of soil S1 = 1, S2 = .01, and S3 = 1, velocity updating parameters V1 = 1, V2 = .01, and V3 = 1. And at last global and local constants βL and βG are initialize by zero point nine values leading to constant values may lead to slight changes as per propose method requirement.

Generate population

The term population has defined to be cumulative accumulation of chromosomes, amongst that very single chromosomes is said by accumulations of predicted pages as per Right Hand Side Pages of association rule. Therefore for the referring chromosome which contains collection of pages tend to behave as predicted pages. So to have a clear view on these accepts assume that dataset have {P7, P2, P11, P12} number of rules having LHS P1, P10 so one of possible solution is Chromo = {P7, P2}. In the same fashion of another group of chromo = {P7, P12}.

PC←Generate Population (Rules)

Selection Probability

Now we can take as N-1 number of drops, here the abbreviation SP is coined to be probability of selection for ever motility[14]. Evaluation for these value is be determined using the given equation:

$$SP(i,j) = \frac{FS(i,j)}{\sum_{k=1}^{N} FS(i,k)}$$

$$FS(i,j) = \frac{1}{\delta + GS(i,k)}$$

$$GS(i,j) = \begin{cases} soil(i,j) & \text{if } \min(soil(i, \text{all element}) > 0 \\ soil(i,j) - \min(soil(i, \text{all element}) & \text{otherwise} \end{cases}$$

Streamlining the Velocity

Velocity can be streamlined by using the i^{th} drop going towards jth node by using the below equation:

$$V(t+1) = V(t) + \frac{V_1}{V_2 + V_3 * Soil(i,j)^2}$$

Revising the value of Soil

Streamlined velocity of the i^{th} drop going toward j^{th} node by using the below equation:

$$\Delta S(i,j) = \frac{S_1}{S_2 + S_3 * T(t+1)^2}$$

$$T(t+1) = \frac{HD}{V(t+1)}$$

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

HD is a term coined to as a heuristic durability, its nothing but a constant value which can lie between 0-1.

$$Soil(i,j) = (1-\beta_L) * Soil(i,j) - \beta_L * \Delta S(i,j)$$

Fitness Function

Fitness of each chromosome were obtained by passing the RInput as per the chromosome pages. Value of X1 and X2 for Chromo = {P7, P2} is obtain by:

$$R_{Input} = \begin{matrix} X_1 & X_2 & X_3 \\ C_{rule1x}\{P_1, P_8\} & C_{rule1x}\{P_1, P_2\} & C_{rule1x}\{P_8, P_2\} \\ C_{rule2x}\{P_1, P_8\} & C_{rule2x}\{P_1, P_7\} & C_{rule2x}\{P_8, P_7\} \end{matrix}$$

$$M = \frac{1}{1 + e^{-(\beta_o + X_1 \times \beta_1 + X_2 \times \beta_2 + X_3 \times \beta_3)}}$$

$$F = \sum_{i=1}^{c} M_i$$

Summation of each page probability value act as fitness value for the chromosome. Hence higher fitness value is consider as good solution.

3.4 Global Crossover

The new chromosomes resulted from the crossover method were next analyzed to add-on the latest population. Then this process is carried out by evaluating the fitness value of the new chromosome in case the output is fine when compared to previous chromosome within the cluster with respect to replacement. In this situation by considering the global bests, the new chromosome replaces the corresponding parent in case the expected fitness value turns to be better than the latest chromosome. It has been found that group-level crossover operations perform better when validated against to global crossovers.

Population Updating

The new chromosomes got from the crossover method were again analyzed to incorporate the latently obtained population. Then this survey is carried out by evaluating the fitness value for the obtained latest chromosome in case the expected fitness value turns to be better than the latest chromosome in the similar cluster against replacement. In the case of a global cross, the current new chromosome replaces the correlated superset if the value of fitness is finer than the former chromosome. It has been achieved that the group-level crossover method works better when contrasted to the global crossover.

Final Solution

The current work says that after satisfactory steps of iteration best suitable chromosome retained and among those extracted group of pages are said to be the recommended pages for suggested model of the genetic algorithm

4. EXPERIMENT AND RESULTS

4.1 Experiment Setup

Experiment was perform on MATLAB tool installed on system having 4 Giga Bytes of Ram with I3 processor. Comparison of proposed work WPPIWD was done with GFFAR [] and PASO [12].

4.2 Evaluation Parameter:

Precision: A transaction's precision is calculated as the ratio of the number of web pages correctly anticipated to the total number of web pages predicted.

Precision=Estimateded_Exact_pages/Allforecasts

Coverage: Is coined to be the ratio of the number of web pages correctly forecasted to the total number

of web pages accessed by the user is used to calculate coverage.

Coverage=Estimated_Exact_pages/All_Visited_Pages

M-metric is defined in this way and is used with the goal of obtaining a single evaluation measure.

Precision = (2xPrecisonxCoverage) / (Precision + Coverage)

Execution Time: Total time required for the algorithm to run in order to anticipate the page based on various dataset sizes

Data Sets: This work consider real dataset form projecttunel.com website it has provide a weblog of April 2019. This has weblog have 20000 session of 6240 users. Number of pages are 278.

4.3 Result

Table 1: Precision value comparison of Page recommendation algorithms.

Testing Dataset Size Percentage	*WPPIWD*	*GFFAR*	*PASO*
20	0.8438	0.5521	0.3931
30	0.8182	0.5664	0.4233
40	0.7801	0.5594	0.4128
50	0.7227	0.4916	0.3877

Table 2: Coverage value comparison of Page recommendation algorithms.

Testing Dataset Size Percentage	*WPPIWD*	*GFFAR*	*PASO*
20	0.4263	0.3458	0.5
30	0.4120	0.1152	0.1022
40	0.3921	0.1138	0.0912
50	0.3629	0.0919	0.0901

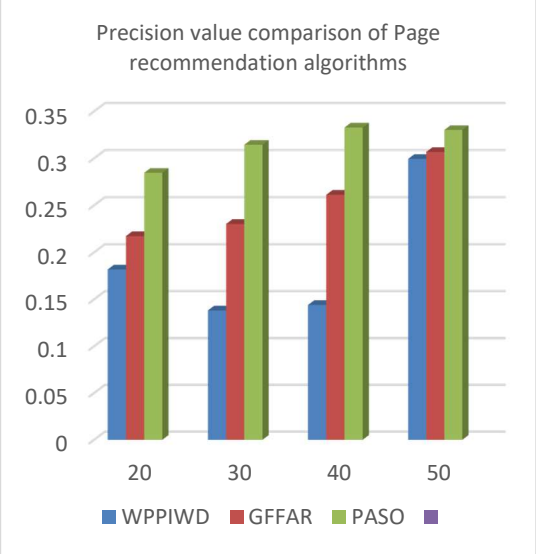

Figure 3: Precision value comparison of Page recommendation algorithms.

Above table 1 and figure 3 shows that proposed WPPIWD has further enhance the precision value towards 1. WPPIWD has improved precision by 31.44% as compared to GFFAR algorithm because of regression model involvement in the fitness function of genetic algorithm. While WPPIWD has improved precision by 48.9% as compared to PASO algorithm. This page prediction efficiency has increase the website performance for user retention.

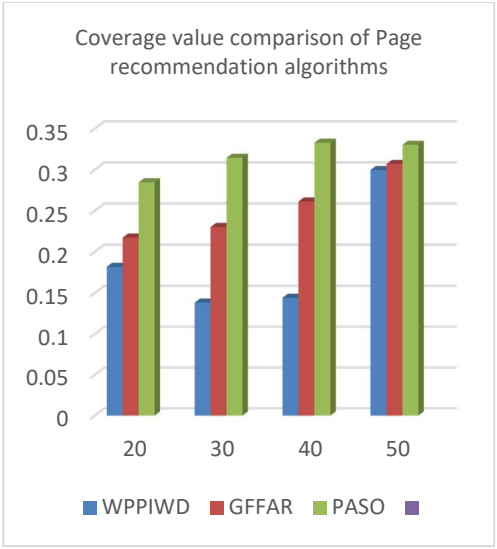

Figure 4: Coverage Value Comparison of Page Prediction Algorithms

Above table 2 and figure 4 shows that coverage value of the proposed work WPPIWD was high as compared to PASO by 58.15% and GFFAR by 50.82%. This enhancement in coverage value was obtained by using the IWD algorithm where algorithm steps has increase the chromosome page set shuffling. Here algorithm has utilized the trained multi regression model for selection of best chromosome where association rule and URL pattern increase its working efficiency.

Table 3: M-metric value comparison of Page recommendation algorithms.

Testing Dataset Size Percentage	*WPPIWD*	*GFFAR*	*PASO*
20	0.1818	0.2173	0.2846
30	0.1382	0.2302	0.3145
40	0.1441	0.2613	0.3328
50	0.2995	0.3068	0.3302

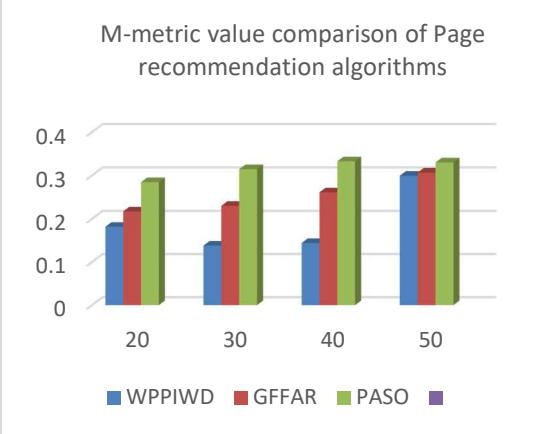

Figure 5: M-metric value comparison of Page recommendation algorithms.

Above table 3 and figure 5 shows that proposed WPPIWD has further enhance the M-metric value towards 1. WPPIWD has improved precision by 49.97% as compared to GFFAR algorithm because of regression model involvement in the fitness function of genetic algorithm. While WPPIWD has improved precision by 57.62% as compared to PASO algorithm. This page prediction efficiency has increase the website performance for user retention.

Table 4: Mean Average Error value comparison of Page recommendation algorithms.

Testing Dataset Size Percentage	*WPPIWD*	*GFFAR*	*PASO*
20	0.5664	0.5248	0.438
30	0.5480	0.1915	0.1646
40	0.5219	0.1891	0.1494
50	0.4831	0.1549	0.1462

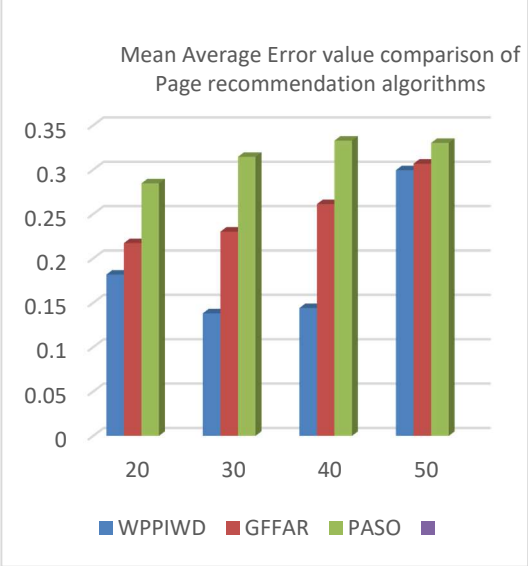

Figure 6: Mean Average Error value comparison of Page recommendation algorithms.

Above table 4 and figure 6 shows that MAE value of the proposed work WPPIWD was reduce as compared to PASO by 24.81% and GFFAR by 39.49%. This reduction in MAE value was obtained by using the IWD algorithm where algorithm steps has increase the chromosome page set shuffling. Here algorithm has utilized the trained multi regression model for selection of best chromosome where association rule and URL pattern increase its working efficiency.

5. CONCLUSIONS

This paper has proposed a new model using machine learning approach of multi regression called frog intelligent water drop model. As an output the model gives the probability of pages

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

predicted in a sequence manner, thus increasing the precision values of predicted pages and also it increases the chromosome page quality involving a minimal value of fitness function with the regression value. The proposed work has concluded with having increasing precision value with minimal fitness value of a precision value when to compare with existing methods like Genetic feed forward association and Particle swam optimization. For improving the performance IWD algorithm was used during testing as this has increase the chromosome page quality involving fitness function with regression values. Live website weblogs were used for the experiment. Results shows that proposed work has increased the precision value by 40.17% as compared to existing methods. In future researcher can improve work accuracy by involving content feature as well. Hence, it deduces that our proposed IWD performs better than current methods. In future, the proposed web page prediction system can well be extended by utilizing additional machine learning algorithms for improving the web page prediction system's performance and for lessening the execution time.

REFERENCES:

[1] J. Dom_enech, J. Sahuquillo, J. A. Gil & A. Pont. The Impact of the Web Prefetching Architecture on the Limits of Reducing User's Perceived Latency. Proc. of the International Conference on Web Intelligence, 2006.

[2] Zukerman, D. W. Albrecht & A. E. Nicholson. "Predicting user's requests on the WWW". Proc. of the seventh international conference on User modeling, pages 275{284, 1999.}

[3] Balamash, M. Krunz& P. Nain. Performance analysis of a client-side caching/pre-fetching system for Web traffic. Computer Network. vol. 51, no. 13, pages 3673{3692, 2007.}

[4] T. M. Kroeger, D.E. Long & J. C. Mogul. Exploring the Bounds of Web Latency Reduction from Caching and Pre-fetching. Proc. of the 1st USENIX Symposium on Internet Technologies and Systems, 1997.

[5] L. Fan, P. Cao, W. Lin & Q. Jacobson. Web Pre-fetching Between Low-Bandwidth Clients and Proxies: Potential and Performance. Proc. of the ACM SIGMETRICS Conference on Measurement and Modeling of Computer Systems, pages 178{187, 1999.

[6] An Ontology-based Webpage Classification Approach for the Knowledge Grid Environment by Hai Dong, FarookhHussain and Elizabeth Chang, 2009 Fifth International Conference on Semantics, Knowledge and Grid (IEEE-2009).

[7] Chong Wang, Shuai Zhao, Achir Kalra, Cristian Borcea, Yi Chen. "Webpage Depth Viewability Prediction Using Deep Sequential Neural Networks Chong Wang ; Shuai Zhao ; Achir Kalra ; Cristian Borcea". IEEE Transactions on Knowledge and Data Engineering Year: 2019 | Volume: 31, Issue: 3 | Journal Article

[8] D.A. Adeniyi, Z. Wei, Y. Yongquan. January 2016. Automated web usage data mining and recommendation system using K-Nearest Neighbor (KNN) classification method. Applied Computing and Informatics Volume 12, Issue 1, Elsevier.

[9] Rahul Katarya, Om Prakash Verma. 2017. An effective web page recommender system with fuzzy c-mean clustering. Multimedia Tools and

[10] Applications, Volume 76, Issue 20, pp 21481–21496, Science+Business Media New York, Springer.

[11] Erman Sen, Hakki Toroslu, Pinar Karagoz. August 2016. Improving the prediction of page access by using semantically enhanced clustering. Journal of Intelligent Information Systems, Volume 47, Issue 1, pp 165–192, Science+Business Media New York, Springer.

[12] T. Arce, P. E. Roman,´ J. D. Velasquez,´ and V. Parada, Identifying web sessions with simulated annealing, Expert Syst. Appl., vol. 41, no. 4 , pp. 1593–1600, 2014.

[13] R. Manikandan. "A novel approach on Particle Agent Swarm Optimization (PASO) in semantic mining for web page recommender system of multimedia data: a health care perspective". Springer Science+Business Media, LLC, part of Springer Nature 10 January 2019.

[14] Nguyen TTS, Lu HY, Lu J (2014) Web-page recommendation based on web usage and domain knowledge. IEEE Trans Knowl Data Eng 26(10):2574–2587.

[15] Rani M, Nayak R, Vyas OP. "An ontology-based adaptive personalized e-learning system, assisted by software agents on cloud storage". Knowl-Based (2015) Syst 90:33–48.