INTERNATIONAL JOURNAL OF INNOVATIVE COMPUTING
ISSN 2180-4370
Journal Homepage : https://ijic.utm.my/

User Behavior for Neural Network-based Web Search Results Filtering

Essam Natsheh
Dept. of Informatics Engineering, College of Engineering,
AMA International University,
Salmabad, Kingdom of Bahrain
Email: dr_natsheh@hotmail.com

Submitted: 21/05/2020. Revised edition: 30/08/2020. Accepted: 1/09/2020. Published online: 19/11/2020
DOI: https://doi.org/10.11113/ijic.v10n2.260

***Abstract*—This paper describes methodology and performance of an experimental research on filtering of web search results. Filtering was performed on the basis of predicted relevance of search results derived from users' implicit feedback. The feedback was obtained from users' web browsers and consisted of a set of browsing behavioral metrics, including reading time, clicks on links, mouse pointer and wheel movement patterns, bookmarking, sharing, copying, and whether the search was continued after the page was closed. A multi-layer neural network used to infer from the behaviors how much the user was interested in each filtered document. Neural network, therefore, performed deep learning without human supervision. Predicted relevance measure was compared to the explicit feedback. Obtained results of 89% correct relevance rating prediction suggest that selected set of metrics was successful in terms of correctly predict how relevant the web page was for the user involved in the study. More research is recommended for further advances of information filtering methods.**

***Keywords*—Search results filtering, browsing behavior analysis, machine learning, reinforcement learning**

I. INTRODUCTION

The amount of generated information has been growing exponentially. Over the past fifty years humanity has generated more than 70% of all information ever created. Although availability and ease of access to knowledge has never been so good, the amount of information in itself poses a challenge to modern people. Even before the age of technology the problem was known under a name of "information overload." The term was initially coined by Bertram Gross, a social scientist, as far back as in the 1964 [1]. By early seventies it became widely known due to the Alvin Toffler's bestseller "Future Shock." Information overload may be defined as the state that occurs when a person making a decision faces far too many alternative pieces (or sources) of information, thus being detrimental to the decision-making process and promoting sub-optimal decisions.

As lion's share of information is sourced from the Web today, and search on the Web using search engines is the primary way of looking for and retrieving the information, the problem of search optimization arises [2]. This optimization could be done by filtering: returning fewer results in an attempt to select the ones more relevant to the user. Although modern search engines utilize various algorithms and approaches for search optimization in an attempt of returning the most relevant results, the outcomes are far from perfect and not tailored personally to the users; that said, the search results vary to some extent between users, e.g., based on their geolocation.

To ensure a truly personalized filtering of Web search results, such service would need to predict a particular user's reaction to displayed Web pages. The reaction in question would be a subjective opinion of the user based on whether the Web page contains the needed information, and whether it is relevant, up to date, concise, correct, and accessible. One way such system could work is through learning based on the user's manual feedback on search results, with continuous reinforcement learning. An immediately apparent downside of this approach would be a significant efforts needed from each user to rate the search results, which are comparable with the need to deal with less relevant results. Therefore, a more useful method would be automatic assessment of how fit each viewed page was to the initial search (as per this particular user's perception). To create such assessment the method could analyze browsing behaviors, including mouse movements, clicks on links and banners, total time spent on the page, whether the search continued, whether any information was copied from the page, whether the page was bookmarked or shared in social media. The recorded metrics could be used as input data for machine learning algorithms.

In this paper, an analysis of the significant factors affecting user experience with search engines and identified the benefits and priority value for these factors on a survey done on search engines users of different profiles between 18-24 years old. The study ascertains that the challenge of indexing and crawling within the internet and general Web users should not be confident that the web search is complete and relevant to what they search for the commercial background of the search service. The economic priority of need for Web-related businesses calls for a higher and complex rank on Web snippets' suggestions to get additional customers. This paper presents a review and analysis of how Neural Network (NN) in Web searches filtering works.

Hypotheses of this study are that NN learning can improve search results relevance by removing part of the results, and that browsing behavior metrics are more practical than direct feedback from users in analyzing the fitness of search results. It is important to note that scope of the study limits the search results for the study to Web pages with text contents, ignoring other media results: video, pictures, etc.

The significance of this research lies in enhancing the process of retrieval of information from the Web, by providing more relevant result with less time spent on the search itself.

II. Literature review

Over the years, filtering based on either explicit or implicit (or both) feedback has been covered in many studies. Our previous papers [3], [4] focused on more recent works in this field.

In [5] researches describe how both implicit and explicit feedback could be utilized simultaneously for the best utilization of available data with the aim of personalized internet information ranking. Their proposed algorithm was based on a combination of the Expected Reciprocal Rank evaluation metric with SVD++ algorithm and resulted in a significant increase in personalized ranking performance.

Patent [6] involves many browsing behavior metrics utilized in search results ranking, including mouse movement (speed, direction and consistency), delay of the user response, and even movement of the user captured by the camera. Some of these parameters could be adopted for the current research.

A study in [7] explores explicit user feedback for evaluation and ranking of search results. They also measure dwell time and discuss limitations of both explicit and implicit methods. What surprised them is a striking difference in dwell time between relevant and irrelevant results that was 87 seconds; this number is much higher than what usually could be found in the literature.

In [8] the relationship between implicit parameters and explicit feedback is investigated. The authors discuss different parameters, such as reading time, mouse movements (distance, clicks), mouse scroll wheel movement, copying of information, highlighting, printing, emailing and bookmarking.

Another approach to enhancing search on the web is interactive intent modeling described in [9]. Its aim is to eliminate the "vocabulary mismatch problem" that occurs between writers and readers by allowing the users to perform search using the user's intended meaning rather than plain words. Albeit this idea as a whole goes beyond this particular research, methods for the interactive intent model adaptation are similar to the ones used for this paper, as they rely on "limited, possibly suboptimal, user feedback." Also, future research on search results filtering might incorporate the idea of intent to better determine relevance of Web pages.

III. Methodology

An experiment was conduct to learn user's profile by observing user behavior. This experimentation was aimed for normal Internet user using web search engine such as Google, Yahoo, etc.

A. Personalized Information Retrieval on the Web (PIRW)

Step 1: Initially, the user provides the system with a profile.

Typically, the initial user's profile (p) consists of one or more keywords. It can be represented as a vector of terms:

$$Wp = (w_{p1},\dots,w_{pk},\dots,w_{pn}) \text{ ,}$$

where

w_{pk}: the weight of k-th term.

n: the number of terms used for describing the profiles.

Fig. 1 shows an example of an initial user's profile and its initial weights.

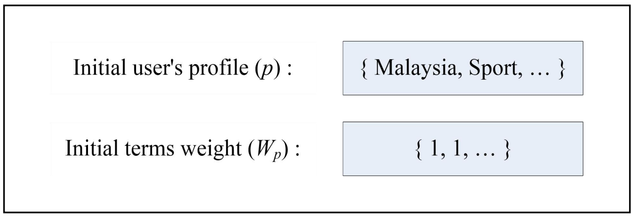

Fig. 1. Initial user's profile and its initial weights

Step 2: (Retrieval) Generate a query from the profile to retrieve N URLs.

Step 3: (Filtering) Evaluate the relevance of documents.

Rank the N documents and present M of them to the user in a descending order (Fig. 2).

The similarity (or relevance) of i document and profile p is computed as:

$$V_{pi} = \sum_{k1}^{n} tf_{ik} \times w_{pk}$$

where

tf_{ik}: term frequency of k-th term in retrieved document D_i

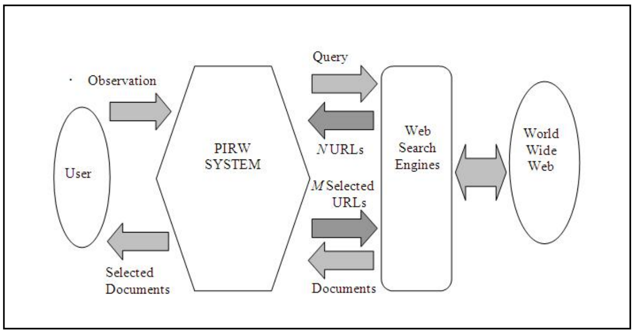

Fig. 2. PIRW will rank the *N* documents and present *M* of them to the user

Step 4: Get the feedback by observing user behavior.

Software in Java language developed to observe the user behavior on the web and called *WebObserver* as shown in Fig. 3. It record following information for each document browsed by the user:

1. Document identifier *(Doc-ID)*.
2. Time for reading *(rt)* the document.
3. Bookmarking *(bm)*.
4. Scrolling thumb up and down *(sl)*.
5. Following up *(fl)* the hyperlinks.

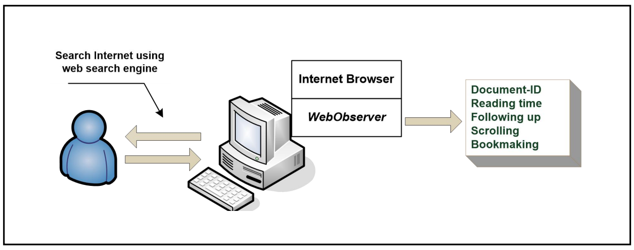

Fig. 3. Get the feedback by observing user behavior

Step 5: (Learning) Update the user profile.

Neural Network (NN) can model a series of complex non-linear relationships that can approximate any measurable function. NN has features that make them a unique and attractive tool for solving patterns recognition tasks least of building an explicit model of the system [10], [11]. In NN modeling in this research, we ascertain architecture and a method of determining the interconnections and transformation function. In the current work, we have identified factors that contribute to increasing user experience and the proposed NN model for evaluating it. NN has been seconded to be an effective tool for web search classification and filtering. The main benefit of NN is that it learns by example and contains the ability to generalize data to other types of data [12]. The search engines normally search for data in their databases using exact words in the query. Since many users are not able to articulate the search with cautious keywords, the search engine fails to generate the expected results or output.

Basically, NN learning consists of inserting new terms, removing existing terms, and adjusting term weights of profile terms using the terms in the relevant/irrelevant documents as shown in Fig. 4. That is, if a term in the document estimated as "relevant" is included in the user's profile, the term receives user's feedback as follows:

$$w_{pk} = w_{pk} + \alpha\, r_i, \qquad \text{if } k \in D_i$$

where

r_i : relevance feedback to the filtered document D_i

α : is the learning rate that controls learning speed.

If a weight of a term is less than the threshold, the term is removed from the user's profile. As the more relevant documents are filtered by a term, the importance of that term will increase.

Neural network can generalize. Immediately when NN learn from the previous inputs and the relationships; they can recognize the hidden relationships data to notable data hence making data model generalized and also predict the hidden dataset. The benefit of NN has a machine learning entity is that it not only creates a network path but also generates a unique code that is disseminated during learning. Its self-organization principle facilitates a clear machine learning experience that fosters a better and coordinated representation of data and information [13]. In real-time operation, the NN calculations can be imposed on a Web Server, which can make the functionality of the proposed system better [14].

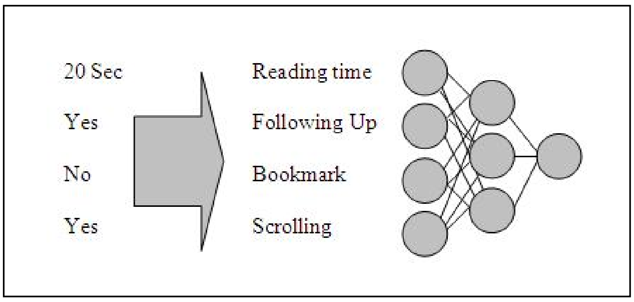

Fig. 4. Neural network used to update the importance of the profile terms

B. Proposed Method Validity

The validity of a measurement technique is the extent to which the technique measures what it is supposed to measure. As our method measure the user's preferences, the participants were asked to validate it after the web browsing sessions were over. They have been asked to rate each document they read in the last session using a developed *ChkDocument* software. This software uses the record of document identifiers (*Doc-ID*) to display contents of each document on users' terminal and interviews the participants about impression of the document as shown in Fig. 5. Then it merges the records described above and the impressions about documents, and sends them to a data collecting server.

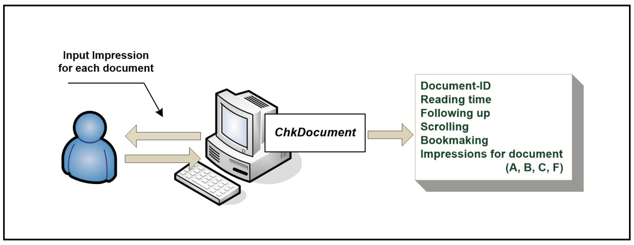

Fig. 5. A method to validate the proposed system

The impressions of document was to be answered in four grade of 'A', 'B', 'C' or 'F' where 'A' meaning *very interesting*, 'B' meaning *interesting*, 'C' meaning *less interesting*, and 'F' meaning *not-interesting*.

IV. PERFORMANCE ANALYSIS

Having designed a Web search results filtering system, it is important to validate its performance. Section *A* describes the implementation technique for the proposed method. The results of this method are presented in section *B* followed by an explanation of key findings.

A. Performance Study

The metrics used for page relevance included:

- Total reading time in seconds (double)
- Number of links clicked on the website (integer)
- Mouse cursor and mouse wheel movements patterns
- Information copied from the page (true or false)
- Bookmarking or social sharing was done (true or false)
- Search continued (another website opened) after the website (true or false)

The above metrics were supplied to a multi-layer NN as input values, and the output was the predicted relevance of the web page from 0 to 10.

The experiment encompassed 100 volunteers aged from 18 to 24. This age group used is mostly college students, and their user experienced was researched. 73 of the participants were males and 27 – females. Users were instructed to interact with the browser and search engine as close to their natural way of browsing as possible.

The users were asked to find answers to three sets of questions on the Web using a search engine. Each set contained 20 questions to answer, similar for all participants. Questions were ranging from simple queries (e.g., "what is the height of Everest"), to broader, often without a "correct" answer, such as "explain how and why Marxism evolved" and "how to bake a good ciabatta." The users were the judges of how well a certain page helped them in answering the question. Therefore, correctness of their answers was not checked.

To test model fitness the users were asked to express their perceived feedback on the relevance of web pages that were returned by the search engine and filtered by the machine learning algorithm. By comparing the predicted relevance value to explicit user's feedback, the model was allowed to train iteratively and reduce loss. Deviation of predicted value as compared to the explicit feedback of either 1 or -1 was deemed a success. Loss was calculated as a difference between the explicit feedback and prediction based on implicit observations.

B. Results

Each of the 100 participants completed three sets of 20 questions, therefore, in total 6000 questions were answered. 7481 search queries were made by the participants. 25 659 web pages were opened by the participants and rated by the algorithm implicitly and by the users explicitly. Average amount of web pages opened per search query was 3.43. Average reading time for a web page was 19.33 seconds. Average amount of links clicked on opened web pages was 0.63 (16 165 clicks in total).

After the initial training, the NN was able to implicitly predict the relevance rating value (1 to 10) within allowed deviation range (from negative 1 to 1) in 83% of the cases. By the end of the trial due to the effects of iterative training approach the prediction quality has improved to 89%.

The results show that the NN filter program starts a loop with each logged packet in which a new evaluation is performed, and it retrains the neural network, preferably according to the known pattern, to adjust the weights assigned to each node to produce an evaluation of the selected packet. Numerous iterations of it are performed, with weights being adjusted until the output rating converges with the user - assigned error threshold. The preferred embodiment offers an alternative to current filtering software techniques that filter inappropriate content or search for words only. In his research presented at Knowledge-based training, the study used neural networks to orient Web search engines towards optimizing search results. It used them to train a wide range of users browsing behaviors. By training the neural network on its wide-range user's behaviors, it can assign accurate evaluations to various documents browsed by the user that fall within the scope of the trained document package. Users browsing behaviors are considered by assigning them an evaluation, similar to ratings for the packages with the nearest input documents used for the training.

C. Discussion

The most impactful behavioral metrics for this model seemed to be whether the search continued, information copied from the page, mouse movement patterns, and to some extent reading time and links clicked on the website. Reading time does not always reflect the relevance of information, but there is a positive correlation to the implicit relevance rating. The amount of links clicked on the website does not always come into play, but there is still a significant positive correlation. Very strong negative correlation was observed to whether the search continued, and strong positive – to whether the information was copied from the page. These results are partly due to the nature of the task – finding answers to predefined questions. Outside of the test scenarios, the author predicts more weight in the bookmarking and social sharing.

The obtained values of prediction accuracy suggest that the neural network is successful in assessing the relevance of search results. Further, the observed increase in accuracy after additional training time could be attributed to the usefulness of continued iterative machine learning in this application. The obtained relevance rating could be used in filtering the search results based on a defined threshold value.

V. CONCLUSIONS

This research explored possibility of utilization of a machine learning method based on implicit user feedback obtained from browsing behavioral patterns. The paper suggested a novel combination of behavioral metrics as input values for a multi-layered neural network. For iterative training of the neural network users' explicit feedback was used. After the experiment was conducted fitness of the obtained filtering model was as high as 89%, which suggest successful implementation of search results relevance rating prediction. These results may be too optimistic due to the inherent differences between experimentation procedure and real-life browsing and searching on the Web. Therefore, the author argues that continued research of the presented methods is needed to confirm the results of this paper. Further research might include introducing variations to the set of metrics to possibly find a better combination of them.

ACKNOWLEDGMENT

The author would like to express his appreciation to College of Engineering at AMA International University Bahrain for supporting this research.

REFERENCES

[1] F. A. Batarseh and R.Yang. (2020). *Data Democracy: At the Nexus of Artificial Intelligence, Software Development, and Knowledge Engineering*. 1st Edition, Academic Press.

[2] C. Smith and S. Rieh. (2019). Knowledge-context in Search Systems: Toward Information-Literate Actions. *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval*, 55-62.

[3] E. Natsheh. (2013). Personalized Web Documents Filtering by Analyzing User Browsing Behaviors. *International Journal of Information Studies (IJIS)*, 5(2), 57-65.

[4] E. Natsheh. (2019). Performance of Standard and Customized Search Interfaces. *International Journal of Human and Technology Interaction (IJHaTI)*, 3(2), 17-21.

[5] G. Li and Q. Chen. 2016. Exploiting Explicit and Implicit Feedback for Personalized Ranking. *Mathematical Problems in Engineering*, 2016, 1-11.

[6] N. Gur. (2016). *U.S. Patent No. 9,449,093*, Washington, DC: U.S. Patent and Trademark Office.

[7] J. Y. Kim, J. Teevan, and N. Craswell. (2016). Explicit In Situ User Feedback for Web Search Results. *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 829-832.

[8] S. Akuma, R. Iqbal, C. Jayne, and F. Doctor. (2016). Comparative Analysis of Relevance Feedback Methods Based on Two User Studies. *Computers in Human Behavior*, 60, 138-146.

[9] T. Ruotsalo, G. Jacucci, P. Myllymäki, and S. Kaski. (2015). Interactive Intent Modeling: Information Discovery Beyond Search. *Commun. ACM*, 58(1), 86-92.

[10] S. Monemian. (2020). A Neuroevolutionary Neural Network-Based Collaborative Filtering Recommendation System. Diss. Laurentian University of Sudbury.

[11] H. Khatter and A. K. Ahlawat. (2020). An Intelligent Personalized Web Blog Searching Technique Using Fuzzy-based Feedback Recurrent Neural Network. *Soft Computing*, 1-13.

[12] J. Kim, and K. Chung. (2020). Neural-network Based Adaptive Context Prediction Model for Ambient Intelligence. *Journal of Ambient Intelligence and Humanized Computing*, 11(4), 1451-1458.

[13] H. Said-Ahmed and E. Natsheh. (2020). Efficient Signatures Verification System Based on Artificial Neural Networks. *International Journal of Human and Technology Interaction (IJHaTI)*, 4(1), 1-9.

[14] E. Natsheh. (2018). The Efficient Use of a List of Trusted Certificate Authorities. *International Journal of Engineering Technology and Sciences (IJETS)*, 5(3), 118-131.