www.gphjournal.org
Gph-International Journal Computer Science Engineering

10.5281/zenodo.14856761 Volume 8, Issue 1 (Jan-Jun 2025) Manuscript ID: #1764

An Improved Model for the Evaluation of Search Engine Result Pages Using Hybridized Evaluation Techniques

P. N. Idyorough
Department of Computer Science and Informatics,Federal University Otuoke, Nigeria.

I. T. Ayorinde
Department of Computer Science, University of Ibadan, Nigeria.

Corresponding author: lizpros7@yahoo.com

Abstract

In this paper, the hybridised ranking system was developed by utilising WPR to evaluate the structural importance of web pages by assigning values to both the incoming and outgoing links and also other web pages it is connected to while incorporating CTR metrics to account for user activities and TF-IDF to reduce the weight of commonly used words by assigning lower values to words that appear in many documents hence combining both to measure content relevance. The hybridised approach is compared with the standard PageRank (PR) using the sum of WPR, CTR and TF-IDF metrics for position ranking with 50 randomly selected query datasets which were ranked into less relevant (LR) and most relevant (MR) positions. The hybridised approach reduced irrelevant pages by reshuffling the less relevant pages to the bottom of the list of web pages found while moving the more relevant pages to the top for the given queries, thereby improving the ranking efficiency of the search engine result pages. In Table 4.1, the process of ranking the 1^{st}, 2^{nd},7^{th} and 8^{th} document for the PageRank is moved to the 23^{rd} 24^{th} 25^{th} and 26^{th} position based on their relevance to the queried data while the 3^{rd}, 4^{th},5^{th} and 6th document maintained a high position based on their relevance. In Table 4.2, the 1^{st} 2^{nd} 3^{rd} and 4^{th} documents were reshuffled to the 7^{th} 8^{th} 9^{th} and 10^{th} positions while the 6^{th} 7^{th} and 8^{th} documentsmaintained a high-ranked position due to their relevance.This study has provided a scalable framework that enhances user experience by minimising irrelevant search results and prioritising pages of higher relevance.

Keywords:

Search engines, ranking algorithm, Weighted PageRank algorithm, Term frequency-inverse document frequency, Click-through rate.

How to cite: Idyorough, P. N., & Ayorinde, I. T. (2025). An Improved Model for the Evaluation of Search Engine Result Pages Using Hybridized Evaluation Techniques. *GPH-International Journal Of Computer Science and Engineering*, 8(01), 01-12. https://doi.org/10.5281/zenodo.14856761

1. INTRODUCTION

The World Wide Web is a massive, extensive and immense knowledge repository, which in itself is a globally distributed information center for knowledge reference and information acquisition. Since its inception, there has been rapid growth doubling the number of resources available [1], these available resources include text, documents, videos, audio and images.

With such a massive data knowledge centre to search for documents, retrieval is usually a huge task involving the use of links which are uniform resource locator (URL) and this is done with the use of search engines. Search engines perform various tasks using different algorithms based on the search engine architecture, Examples of search engines include Google search engine, Yahoo search engine, Bing among others. The basic components of the search engine are the interface, parser, web crawler and the database [2].The interface serves as an interactive point between the user and the system. The web crawler does the searching for the user request from an immense stack of documents stored in a database. The documents are taken and split into indexable text fragments by a parser then the ranking engine ranks the documents with similar keywords that relate to the user requests before returning the results to the user. Further analysis of the activities of the user points to the visits to several pages in search of documents with more relevance to the search which creates an incoming and outgoing link to documents [3]. When a search engine returns the documents found, a ranking algorithm is used to prioritize the documents based on the keywords associated with the user query. Such algorithms developed include PageRank, Hyper-link Induced Topic Search (HITS), SIMRank, Randomized HIT, among others [4]. The algorithms developed analyze the documents using content or links without the consideration of user usage trends [4].For a link analysis algorithm, a given Root set is retrieved using a text-based web search engine which consists of a relatively short list of web pages relevant to a given query, the Root set is improved by pages that point to the pages in the Root set and other pages that are pointed to by pages in the Root set hence obtaining a larger base set of web pages on which the algorithm functions [5] From the base set of pages, the hyperlink is developed from a node which is generated from every web page with a directed edge placed between two nodes as a hyperlink between the related web pages. The graph is made simple even when there are multiple links between pages, only a single edge is placed. The webpages are analyzed using the content of the pages thereby removing isolated pages from the graph.

The PageRank algorithm is believed to be one of the most extensively used page ranking algorithms. It states that when a page has more important links to it, other pages linked to it also become important hence the PageRank recognizes backlinks and utilizes the ranking of the links. A page is said to have a high rank when its backlink pages have higher ranks.

According to [6] this algorithm considers the link structure and not the content of a page hence its ability to include less relevant pages is high, also [7] in their work stated that the page's relevance to a certain query was less determined and the utilisation of the web structure had disregarded some criteria which has the possibility to significantly produce pages with other outcomes.

According to [8] Klienberg gave two forms of webpages called Hubs and Authorities, hubs are the pages that act as the resource list while authorities are the pages with important content, According to [3] HITs ranks works by the analysis of both the in-links and out-links, the webpages pointing to many hyperlinks are referred to as the hubs whereas the webpages being pointed to by many hyperlinks are called the authorities. A page can both be a good hub and a good authority as well, both hubs and authorities are assigned scores respectively. According to [9] the hyperlink information of a given page includes a number of links, anchor text, and positions of the pages in the domain tree concerning a particular page.

The number of hyperlinks: in calculating the number of hyperlinks on a page the number of frame source tags, and the number of href tags are added but the links to the same page are excluded. The Anchor text: in determining the hub and authority weight, the weight of links can be determined using the anchor text, and the glossary pages can be easily recognized and analysed using the anchor page. The position of the pages in the domain tree with respect to a particular page: the portals having a lot of links are connected to the same level nodes in the domain tree which are rooted at the next higher-level node of the page's source[10].

According to[11] as the vast amount of available data increased, semantic ranking gained significant relevance. He created a scenario in which documents were said to be semantically connected to the author's area of expertise where the most pertinent results did not appear when attempting to locate an author associated with a query using standard text similarity matching methods such as TF-IDF, hence in such cases, the application of semantics became necessary.

2. Literature Review

2.1A Weighted PageRank-Based Bug Report Summarization Method Using Bug Report Relationships [12]: Proposed a bug report summarisation method that uses weighted PageRank algorithm. They used the algorithm for sharing and discussing information, checking past changes as well as referring to relevant bug fixes.

2.2 Augmented Graph-based Unsupervised Key phrase Extraction [13]: Proposed an augmented graph-based unsupervised model to identify key phrases from a document by integrating graph and deep learning methods. This model utilizes mutual attention while also evaluating on four datasets.

2.3 Comparative study of various Page RankingAlgorithms in Web Structure Mining (WSM)[14]: Proposed the assignment of more rank value to the outgoing links which are most visited by users and received higher popularity from several in-links nevertheless assigned more rank value to the outgoing links which are most visited by users and received higher popularity from several in-links. This work did not include the TF-IDF algorithm.

2.4 Weighted page rank algorithm based on in-out weight of webpages[15]: they introduced a new weight matrix based on both the in-links and out-links between web pages to compute the page ranks but then used only the in and out links of the webpage.

2.5 Weighted PageRank using the Rank Improvement [16]: Uses the relevancy values for the query produced by Page Rank and Weighted PageRank using different page sets then uses only the in-links and out-links of a page.

2.6. Page Ranking Based on Number of Visits of Links of Web Page [17]: In this paper, a PageRank algorithm is used and takes into cognizance the number of visits of inbound links of Web pages into account. Nonetheless, the click-through rate was not utilised.

2.7 Weighted PageRank algorithm [18]: Introduced in this paper Weighted PageRank, takes into account the importance of both the in-links and the out-links of the pages and distributes rank scores based on the popularity of the pages nonetheless it distributes rank scores based on the popularity of webpages.

3. PURPOSE OF THIS PAPER

The purpose of this paper is to improve search engine result ranking efficiency by incorporating users' search behaviours into the information ranking process that meets users' needs and saves time. This study proposes the optimization method of the associative knowledge graph using TF-IDF-based ranking scores. The proposed method calculates TF-IDF weights in all documents and generates term ranking. Based on the terms with high scores from TF-IDF-based ranking, optimized transactions are generated. This work used the TD-IDF algorithm alone.

This paper is structured as; section1, introduction, section 2,literature Review, section 3 present the material and methods for searched queries, section 4 presents the computational results for searched queries.

4. Materials and Methods

4.1. The Click-Through Rate Technique

The Click-through rate (CTR) is the number of clicks by the user on the returned result page, the user browses through the results and clicks on the link they think is relevant to the query. However, clicks are monitored such that for a link ranking top on a SERP, when most users click on it and bounce right back it is seen that the page does not contain the right content for the user hence the page is demoted and a page that the users click on and spend more time on is moved to the top and given a higher value than that which was at the top. The chart below illustrates the process involved in calculating the click-through rate,

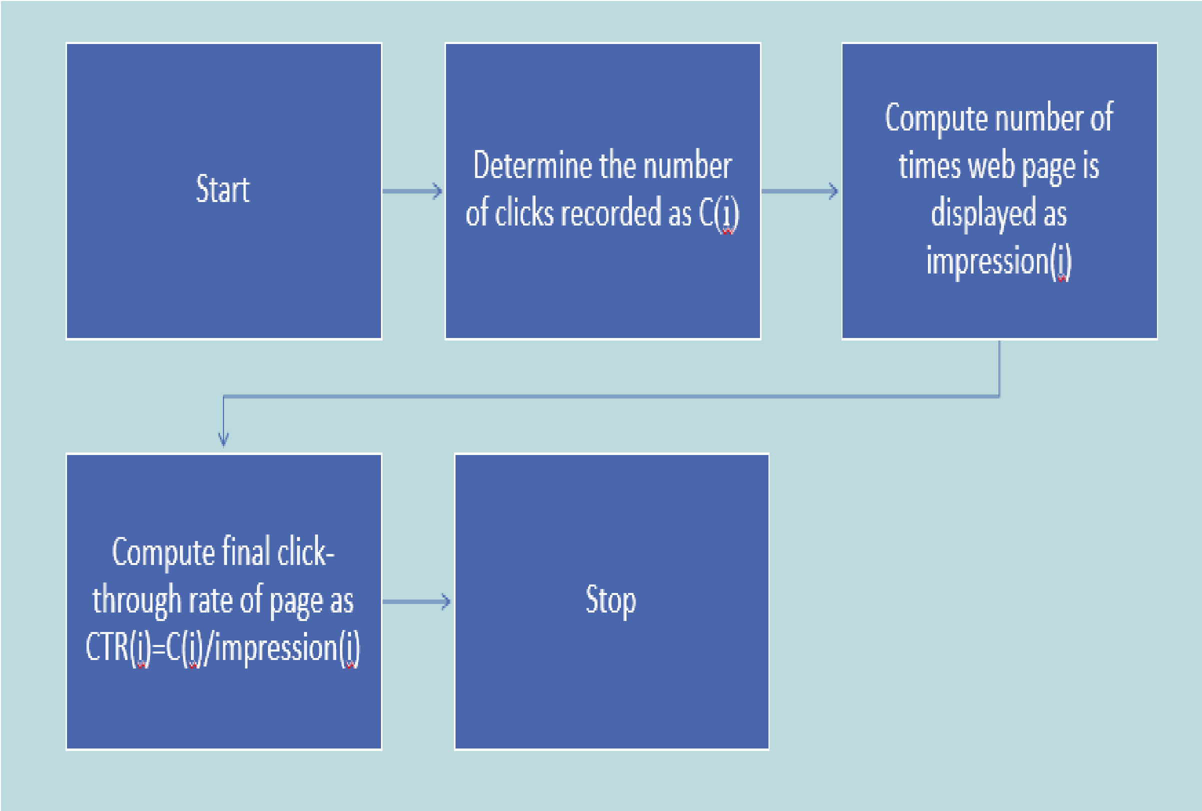

Figure 4.1 Block Process of the Click-Through Rate.

Figure 4.1 above depicts the flow of the click through rate whereby a webpage is analysed from the start position with the number of clicks recorded as $c_{(i)}$ is determined then the number of times a web page is displayed as impression(i) is computed after which the click-through rate of the page is computed as CTR equals c(i) divided by the impression(i) to increase the value of the pages found.

The value of the Click Through is calculated given the formula below;

$$\mathrm{CTR}(d) = \frac{\mathrm{Clicks}(d)}{\mathrm{impressions}(d)} \tag{4.1}$$

Where, $\mathrm{Clicks}(d)$ is the number of clicks that the web page d has received.

$\mathrm{impressions}(d)$ is the number of times the web page d has been displayed in search results.

$\mathrm{CTR}(d)$ is the click-through rate of web page d, calculated as the ratio of clicks to impressions.

By including CTR in the search ranking algorithm, we can adjust the ranking of search results based on how frequently users click on the web Pages. Pages with a high CTR are likely to be more relevant to the user's query and therefore are ranked higher in search results while those without clicks are drawn to the top using the $TF-IDF$ algorithm which has been established to evaluate a webpage and rank based on the relevance of its content.

4.2 Weighted Page Rank

In the weighted PageRank algorithm (WPR), the more web pages are popular the more the linkages that other web pages tend to have to them. The WPR assigns larger rank values to more important pages instead of dividing the rank value evenly among its out-link pages. Each out-link page has a value proportional to its popularity. The diagram below illustrates the workflow of how the values in WPR are generated, where the values of the non-directed link set are calculated, after which the rank score for the page is calculated and weights are assigned to both outbound and inbound links. The value of the dampingfactor, a constant set at 0.85, is multiplied by the value of the summation of the rank score and link weights, and the weight of the PageRank is finally factored in to give the weighted value.

The weighted page rank algorithm is given by the formula;

$$WPR(u)=(1-d)+d\sum_{v\in B(u)}WPR(v)W_{in}(u,v)W_{out}(u,v) \quad (4.2)$$

Where, $WPR(u_1,u_2,...,u_n)$ is the weighted page rank (incoming links and outgoing links) of a page u from a particular search, d is the damping factor, the sum of the weight from different pages is summed and the value of the number of times the page was visited is added.

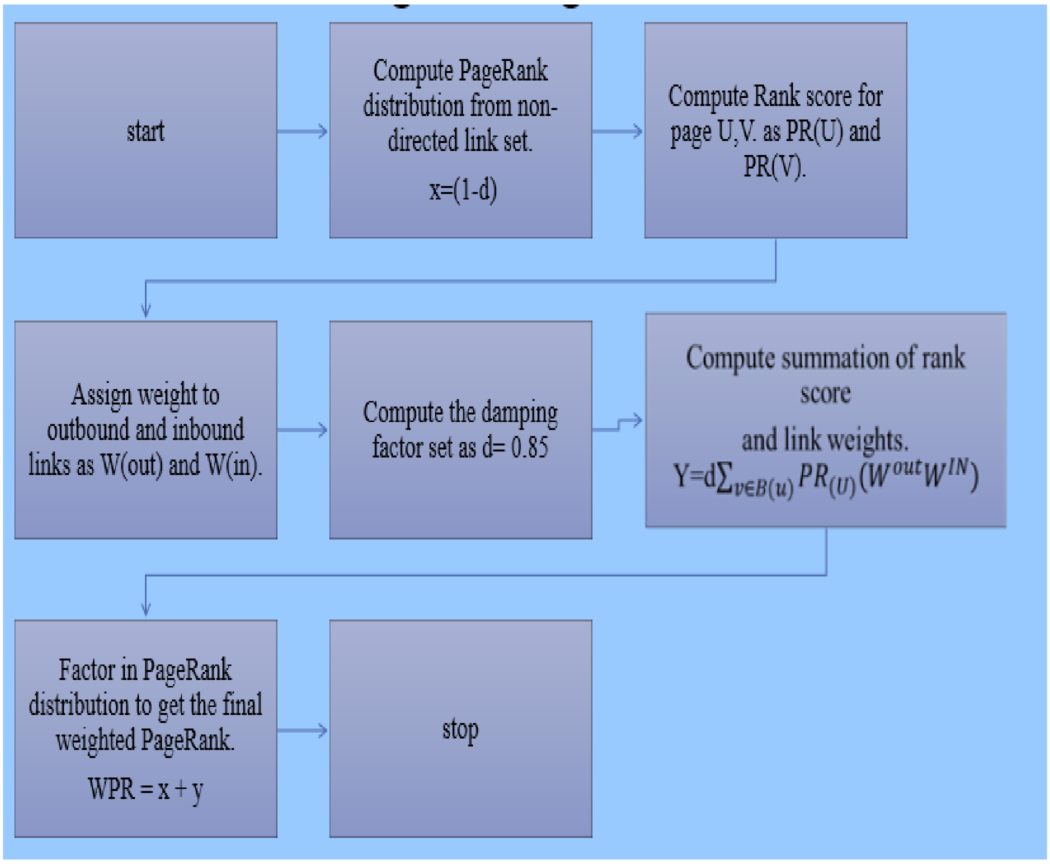

Figure *4.2*: Block Process of WPR.

$WPR : W^{in}_{(v,u)}$ is the weight of link (v,u) calculated based on the number of in-links of page u and the number of in-links of all reference pages of page v.

$$WPR(u) = (1-d) + d \sum_{v \in B(u)} PR(v) W^{out}_{(u,v)} W^{in}_{(u,v)} \quad (4.3)$$

Where, $PR(u)$ and $PR(v)$ are rank scores of web pages u and v, respectively. calculated iteratively using the PageRank algorithm.

$(1-d)$ as the PageRank distribution from non-directly linked pages.

d is the damping factor used in the PageRank algorithm set to 0.85

$W^{in}_{(i)}$ and $W^{out}_{(i)}$ are the weights assigned to the inbound and outbound links of webpage u, respectively.

4.3 Term Frequency- Inverse Document Frequency $(TF - IDF)$

The term frequency- inverse document frequency works by determining the relative frequency of words in a specific document as compared to the inverse proportion of that word over the entire document thereby calculating the relevance of a given word in a document. Common words such as articles and prepositions are usually given lower values than other words making them relatively negligible unless otherwise specified by the user.

This algorithm determines the relevance of a given word in a document thus uses the concepts that TF-IDF assigns to term t a weight in document d that is highest when t has a high occurrence in a small document, lower when the term occurs less frequently and lowest when the term t occurs in virtually all documents. Pronouns and prepositions found in document usually hold no relevant meaning in a query except when specifically included in the search hence given a very low relevant score that is negligible. The structure chart below describes the workings of $TF - IDF$ whereby it fetches the indexed document after they have been crawled from the web to the base set to count the total number of words in a document while also checking the occurrence of each word in all the extracted document then applying the TF formula, afterwards the documents are checked if they contain the query words and the total number of documents is counted then the IDF formula is applied, thereafter the

$TF-IDF$ formula is applied on all the documents to mark the end process and this is done to ensure new documents are ranked better on the search engine result page.

The Term Frequency formula is applied to the document and the total count of documents is taken to apply the inverse document formula, the $TF-IDF$ formula is applied to a new document to calculate its popularity as well as increase its rank value. The equations used in this process are as follows.

$$TF-IDF(t,d)=TF(t,d)\times IDF(t) \quad (4.4)$$

Where, $TF(t,d)$ is the frequency of the term t in document d and

$IDF(t)$ is the inverse document frequency of term t, calculated as the logarithm of the total number of documents N divided by the number of documents containing term $t(n(t))$.

$TF-IDF(t,d)$ is the $TF-IDF$ score of term t in document d, calculated as the product of $TF(t,d)$ and $IDF(t)$..

Search ranking based on human behavior:

$\text{Score}(q,d)$ is the overall score of webpage d for the search query q, calculated as the sum of the TF-IDF and WPR scores of all the terms in q that appear in d. To incorporate the values of all the algorithms into the search ranking algorithm for proper reranking of documents before being displayed, the Score expression is as follows;

$$\text{Score}(q,d)=\sum_{t\in q}\left\{TF-IDF(t,d)+WPR(d)+CTR(d)\right\}$$

(4.5)

Where, q is the search query, represented as a set of terms,

d is the web page that matches the search query,

$TF-IDF(t,d)$ is the score of term t in the webpage d,

$WPR(d)$ is the weighted PageRank score of webpage d,

$CTR(d)$ is the click-through rate of web page d.

$\text{Score}(q,d)$ is the overall score of web page d for the search query q, calculated as the sum of the $TF-IDF$, WPR, and CTR scores of all the terms in q that appear in d.

5. Result

This section presents the computational results for the searched queries.

Table 5.1: Ranked result for the searched term "Data".

Rank	Title	Link	ID	TF-IDF Score	WPR Score	CTR	WPR + CTR	WPR + TF-IDF	TF-IDF + CTR
1	Maps	Link	0	0.0000	0.0000	1.0000	1.00000	0.0000	1.0000
2	English (US)	Link	3	0.0000	0.0343	0.5000	0.53435	0.0343	0.5000
3	Learn more	Link	28	0.0000	0.0343	0.3333	0.36765	0.0343	0.3333
4	Sign in	Link	29	0.0000	0.0343	0.2500	0.28435	0.0343	0.2500
5	Data analysis	Link	4	0.1105	0.0335	0.2000	0.23351	0.1440	0.3105
6	Data acquisition	Link	5	0.1105	0.0335	0.1667	0.20021	0.1440	0.2772
7	Data administrator	Link	6	0.1105	0.0335	0.1429	0.17641	0.1440	0.2534
8	Data (computer science)	Link	9	0.1105	0.0335	0.1250	0.15851	0.1440	0.2355
9	Data (disambiguation)	Link	10	0.1105	0.0335	0.1111	0.14461	0.1440	0.2216
10	Data analysis	Link	11	0.1105	0.0335	0.1000	0.13351	0.1440	0.2105
11	Data science	Link	12	0.1105	0.0335	0.0909	0.12441	0.1440	0.2014
12	Liverpool's former director of research Ian Grahar	Link	13	0.1105	0.0335	0.0833	0.11681	0.1440	0.1938
13	NDPC Fines Fidelity Bank N555.8m over Alleged B	Link	14	0.1105	0.0335	0.0769	0.11041	0.1440	0.1874
14	Oil Minister disputes OPEC data, insists Nigeria's	Link	15	0.1105	0.0335	0.0714	0.10491	0.1440	0.1819
15	Fidelity Bank Fined N555.8m By NDPC For Data Br	Link	16	0.1105	0.0335	0.0667	0.10021	0.1440	0.1772
16	Nigerian govt fines Fidelity Bank N555.8 million o	Link	17	0.1105	0.0335	0.0625	0.09601	0.1440	0.1730
17	FG fines Fidelity Bank N555.8m over data breach	Link	18	0.1105	0.0335	0.0588	0.09231	0.1440	0.1693
18	NDPC fines Fidelity Bank N555.8 million over data	Link	19	0.1105	0.0335	0.0556	0.08911	0.1440	0.1661
19	After Iran Steals Sensitive Israeli Data, Israel Tries t	Link	20	0.1105	0.0335	0.0526	0.08611	0.1440	0.1631
20	Nigerian Fines Fidelity Bank Record N555.8Million	Link	21	0.1105	0.0335	0.0500	0.08351	0.1440	0.1605
21	NDPC fines Fidelity Bank N555m for 'violating' dat	Link	22	0.1105	0.0335	0.0476	0.08111	0.1440	0.1581
22	Data Definition & Types of Sources - Lesson - Stu	Link	23	0.1105	0.0335	0.0455	0.07901	0.1440	0.1560
23	Data Definition & Meaning - Merriam-Webster	Link	1	0.2211	0.0327	0.0435	0.07621	0.2538	0.2646
24	DATA English meaning - Cambridge Dictionary		2	0.2211	0.0327	0.0420	0.07471	0.2538	0.2631
25	What is Data? - Definition from WhatIs.com - Tech	Link	7	0.2211	0.0327	0.0400	0.07271	0.2538	0.2611
26	Data - Wikipedia	Link	8	0.2211	0.0327	0.0385	0.07121	0.2538	0.2596
27	What is Data? Embracing the Basics and Its Impor	Link	25	0.2211	0.0327	0.0370	0.06971	0.2538	0.2581
28	Nigerian govt fines Fidelity Bank N555m for allege	Link	24	0.2211	0.0327	0.0357	0.06841	0.2538	0.2568
29	Fidelity Bank fined N555.8m by NDPC for data bre	Link	26	0.2211	0.0327	0.0345	0.06721	0.2538	0.2556

Table 5.2 Displays the ranking for the searched term "Categorical Data".

Rank	Title	Link	ID	TF-IDF Score	Weighted PageRank Score	CTR	WPR + CTR	WPR + TF-IDF	TF-IDF + CTR
1	Maps	Link	0	0.0000	0.0498	1.0000	1.0498	0.0498	1.0000
2	Learn more	Link	20	0.0000	0.0498	0.5000	0.5498	0.0498	0.5000
3	Sign in	Link	21	0.0000	0.0498	0.3333	0.3832	0.0498	0.3333
4	Categorical Data	Link	7	0.1146	0.0487	0.2500	0.2987	0.1633	0.3646
5	Categorical variable - Wikipedia	Link	8	0.1146	0.0476	0.2000	0.2476	0.1622	0.3146
6	Wikipedia	Link	6	0.0000	0.0476	0.1667	0.2143	0.0476	0.1667
7	Categorical Data: Definition, Types, Features + Examples	Link	1	0.2404	0.0475	0.1429	0.1904	0.2879	0.3833
8	What is categorical data?	Link	2	0.2404	0.0475	0.1250	0.1725	0.2879	0.3654
9	Types of categorical data	Link	3	0.2404	0.0475	0.1111	0.1586	0.2879	0.3515
10	Features of categorical data	Link	4	0.2404	0.0475	0.1000	0.1475	0.2879	0.3404
11	Categorical Data Overview, Analysis & Examples - Lesson	Link	10	0.2404	2.4044	0.0475	2.4520	2.6448	0.2879
12	What is Categorical Data?	Link	11	0.2404	0.0475	0.0833	0.1309	0.2879	0.3237
13	Categorical Data Examples	Link	12	0.2404	0.0475	0.0769	0.1245	0.2879	0.3173
14	Categorical Data Analysis	Link	13	0.2404	0.0475	0.0714	0.1190	0.2879	0.3118
15	Examples of Categorical Data	Link	15	0.2404	0.0475	0.0667	0.1142	0.2879	0.3071
16	Analysis of Categorical Data	Link	16	0.2404	0.0475	0.0625	0.1100	0.2879	0.3029
17	What is Categorical Data? Definition, Types, Examples Appinio Blog	Link	18	0.2404	2.4044	0.0475	2.4520	2.6448	0.2879
18	What is Categorical Data? - Definition & Examples - Lesson	Link	19	0.2404	0.0475	0.0556	0.1031	0.2879	0.2960
19	Categorical Data & Qualitative Data (Definition and Types) - BYJU'S	Link	9	0.3662	0.0464	0.0526	0.0991	0.4126	0.4188
20	Types of Data in Statistics: Numerical vs Categorical Data	Link	17	0.3662	0.0464	0.0500	0.0964	0.4126	0.4162
21	Categorical Data: Definition + [Examples, Variables & Analysis]	Link	5	0.4809	0.0455	0.0476	0.0931	0.5264	0.5285
22	Categorical Data: Definition, Types and Examples - GeeksforGeeks	Link	14	0.4809	0.0455	0.0455	0.0909	0.5264	0.5264

DISCUSSION OF RESULTS

In Tables 4.1 and 4.2 the sum of the $TF-IDF$, Weighted PageRank Score and CTR gives us the new Ranking position. The column with the title ID represents the PageRank algorithm with the positions from the searched queries. In Table 4.1, the process of ranking the 1st, 2nd,7th and 8thdocument for the PageRank is moved to the 23rd24th 25th and 26thposition based on their relevance to the queried data while the 3rd,4th ,5th and 6th document maintained a high position based on their relevance.

In table 4.2, The 1st2nd 3rd and 4thdocumentsreshuffled to the 7th8th9th and 10thposition while the 6th7th and 8thdocuments were able to maintain a high ranked position due to their relevance.

CONCLUSION

In this paper, three different ranking algorithms were used, and the results were derived as follows: (i). There was an improvement in the ranking results as a combination of the three algorithms. (ii). A new search algorithm was developed. (iii). The summed values obtained from the combination of the three algorithms displayed had higher values than the individual algorithms hence a better algorithm.

REFERENCES

Rani, R., & Jain, V. (2013). Weighted PageRank using the Rank Improvement. International Journal of Scientific and Research Publications, ISSN: 2250-3153, 3(7), 3–7.

Ullah, A., Nawi, N. M., Sutoyo, E., Shazad, A., & Naqeeb, S. (2018). Search Engine Optimization Algorithms for Page Ranking : Comparative Study. International Journal of Integrated Engineering, 10 No. 6(December), 19–25. https://doi.org/10.30880/ijie.2018.10.06.003.

Xing, W.& Ghorbani, A. (2004). Weighted PageRank Algorithm.

Tyagi, N. & Sharma, S. (2012). Weighted Page Rank Algorithm Based on Number of Visits of Links of Web Page. 3, 441–446.

Borodin, A., Roberts, G. O. & Rosenthal, J. S. (2005). Link Analysis Ranking: Algorithms, Theory, and Experiments. 5(1), 231–297

Joshi, S. & Goel, S. (2021). Comparative Study of Page Rank and Weighted Page Rank Algorithm. SSRN Electronic Journal, 2929–2937. https://doi.org/10.2139/ssrn.3919359

Shaffi, S. S. &Muthulakshmi, I. (2023). Weighted PageRank Algorithm Search Engine Ranking Model for Web Pages. Intelligent Automation and Soft Computing, 36(1), 183–192. https://doi.org/10.32604/iasc.2023.031494.

Rekha, J. & Purohit, G. N. (2010). Page Ranking Algorithms for Web Mining Page Ranking Algorithms for Web Mining. January. https://doi.org/10.5120/1775-2448.

Mukhopadhyay, D., Biswas, P. & Kim, Y. (2006). A Syntactic Classification based Web Page Ranking Algorithm. 83–92.

Byron, E., Ravi, S., Gibson, D. & Kleinberg, J. (2015). Mining the Web' s Link Structure. March. https://doi.org/10.1109/2.781636.

Amjad, T., Daud, A. &Aljohani, N. R. (2018). Ranking Authors in Academic Social Networks: ASurvey. Library Hi Tech, 36(1), 97–128. https://doi.org/10.1108/LHT-05-2017-0090

Kim, B., Kang, S.& Lee, S. (2019). A Weighted PageRank-Based Bug Report Summarization Method Using Bug Report Relationships. Applied Sciences. 9. 5427. 10.3390/app9245427.

Haoran D. and Xiao L. (2022). AGRank: Augmented Graph-based Unsupervised Keyphrase Extraction. In Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), pages 230–239, Online only. Association for Computational Linguistics.

Neelam, T.and Sharma, S. (2012). Comparative Study of Various Page RankingAlgorithms in Web Structure Mining (WSM)International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-1, Issue-1, June 2012.

Jaganathan, B. &Desikan, K. (2015). Weighted Page Rank Algorithm Based on in-out Weight of Webpages. Indian Journal of Science and Technology. 8. 10.17485/ijst/2015/v8i34/86120.

Rashmi. R. & Vinod, J. (2018).Weighted PageRank using the Rank Improvement; Int J Sci Res Publ 3(7) (ISSN: 2250-3153). http://www.ijsrp.org/research-paper-0713.php?rp=P191481.

Kumar, G.&Duhan, N. Sharma, A. (2011). Page Ranking Based on Number of Visits of Links of Web page. 11-14. 10.1109/ICCCT.2011.6075206.

Xing, W. &Ghorbani, A.A. (2004). Weighted PageRank Algorithm. Proceedings. Second Annual Conference on Communication Networks and Services Research, 2004., 305-314.